Adaptive Control of Transit Operations

U.S. Department
of Transportation

Federal Transit
Administration

Report No. MD-26-7002

November 15, 1995

Office of Technical Assistance and Safety

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof

The United States Government does not endorse manufacturers or products. Trade names appear in the document only because they are essential to the content of the report.

Technical Report Documentation Page

1. Report No. MD-26-7002	2. Government Accession No.	3. Recipient's Catalog No. UMTRIS
4. Title and Subtitle Adaptive Control of Transit Operations		5. Report Date November 15, 1995 6. Performing Organization Code
7. Author(s) Guey-Shii Lin, Ping Liang, Paul Schonfeld, Robert Larson		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Civil Engineering, The University of Maryland College Park, MD 20742		10. Grant or Project No. UMTRIS/TRIS 11. Contract or Grant No. MD-26-7002
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Transit Administration (FTA) 400 7th Street, SW Washington, DC 20590		13. Type of Report and Period Covered Final Report 08/15/1993 ~06/15/1995 14. Sponsoring Agency Code FTA/TRI 30
15. Supplementary Notes		

16. Abstract

Bus priority treatments have been used for decades in various efforts to improve transit services. To improve bus movements along their routes without incurring delays to other traffic, two simulation models for bus dispatching control and for adaptive signal control have been developed in this study.

The bus dispatching control model is developed for evaluating, and eventually implementing strategies along bus routes. Traffic signals may also be controlled to favorably influence the movements of buses.

With the two models, several performance measures are analyzed at individual intersections and on bus routes. Holding and stop-skipping controls are analyzed and optimized based on a specified cost function. Headway-based and schedule-based control strategies are compared by various criteria of interest. Traffic operating costs and bus delay costs are also evaluated through sensitivity analysis of parameters such as bus headways, bus delay costs, and signal timing.

The integration of the two models is tested in a network. Traffic data and the simulated intermediate performance measures are communicated between the two models on a node-by-node basis. One numerical result shows that the average bus delay time can be reduced by up to 55% with bus priority control. With 5 minute bus headways, the combined operating cost for buses and other traffic is reduced by approximately 6% with priority control. The cost saving opportunities decrease as bus headways increase.

17. Key Words Bus priority, Holding control, Stop-skipping control, Headway-based strategy, Schedule-based strategy, Signal preemption, Green extension, Red truncation, Transition cycles, Convergence, Fibonacci search.	18. Document Availability Available to the public through the National Technical Information Service (NTIS) 5285 Port Royal Road Springfield, Virginia 22161 Telephone: (703) 487-4650

19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of pages 121 pages	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Adaptive Control of Transit Operations

BY

Guey-Shii Lin
Ping Liang
Paul Schonfeld
Robert Larson

November 7995

Transportation Studies Center
University of Maryland
College Park, MD 20742

Final Report for Project MD-26-7002
Federal Transit Administration

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in) = 2.5 centimeters (cm)
1 foot (ft) = 30 centimeters (cm)
1 yard (yd ❑ 0.9 meter (m)
1 mile (mi) = 1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2 = 6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2 = 0.09 square meter (m^2)
1 square yard (sq yd, yd^2) = 0.8 square meter (m^2)
1 square mile (sq mi, mi^2) = 2.6 square kilometers (km^2)
1 acre = 0.4 hectares (he) = 4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz) = 28 grams (gr)
1 pound (lb) = .45 kilogram (kg)
1 short ton = 2,000 pounds (Lb) = 0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp) = 5 milliliters (ml)
1 tablespoon (tbsp ❑ 15 milliliters (ml)
1 fluid ounce (fl oz) = 30 milliliters (ml)
1 cup (c) = 0.24 liter (l)
1 pint (pt) = 0.47 liter (l)
1 quart (qt) = 0.96 liter (l)
1 gallon (gal) = 3.8 liters (l)
1 cubic foot (cu ft, ft^3) = 0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3) = 0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] ^{0}F ❑ y ^{0}C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm) = 0.04 inch (in)
1 centimeter (cm) = 0.4 inch (in)
1 meter (m) = 3.3 feet (ft)
1 meter (m) = 1.1 yards (yd)
1 kilometer (km) = 0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2) = 0.16 square inch (sq in, in^2)
1 square meter (m^2) = 1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2) = 0.4 square mile (sq mi, mi^2)
1 hectare (he) = 10,000 square meters (m^2) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gr) = 0.036 ounce (oz)
1 kilogram (kg) = 2.2 pounds (lb)
1 tonne (t) = 1,000 kilograms (kg) = 1.1 short tons

VOLUME (APPROXIMATE)

1 milliliters (ml) ❑ 0.03 fluid ounce (fl oz)
1 liter (1) = 2.1 pints (pt)
1 liter (l) = 1.06 quarts (qt)
1 liter (l) = 0.26 gallon (gal)
1 cubic meter (m^3) = 36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3) = 1.3 cubic yards (cu yd, yd^3)

TEMPERATURE (EXACT)

[(9/5) y + 32] ^{0}C ❑ x ^{0}F

QUICK INCH-CENTIMETER LENGTH CONVERSION

INCHES: 0 1 2 3 4 5 6 7 8 9 10

CENTIMETERS: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 25.40

QUICK FAHRENHEIT-CELSIUS TEMPERATURE CONVERSION

^{0}F	-40°	-22^0	-4°	14°	32°	50°	68°	86°	104°	122°	140°	158°	176°	194°	212°
°C	-40°	-30°	-20°	-10°	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°

For more exact and or other conversion factors, see NBS Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50. SD Catalog No. Cl3 10286.

Adaptive Control of Transit Operations

Executive Summary

This project has developed methods for controlling the movement of transit vehicles along their routes through adaptive control. The basic concept is that adjustments in traffic signals and other controlled variables, based on real-time information, may be used to help transit vehicles move at higher average speeds and better adhere to schedules (including, very importantly, meet connecting transit vehicles at transfer stations).

Signal pre-emption has been considered for a long time as a possible option for reducing delays to transit vehicles at signalized intersections. Some pre-emption, based on very simple logic, could be implemented decades ago with much simpler technology than that available today. The control logic might simply be "if a bus is detected or requests a green phase, then turn green immediately, or after a minimum pre-set red duration." Such logic may be fairly adequate for isolated low-volume intersections. Unfortunately, transit delays and service unreliability problems are far more attributable to closely interrelated and congested intersections. Primitive control logic that automatically and immediately favors transit vehicles may cause severe problems to other traffic at the same intersection, may seriously disrupt a coordinated system of signals at nearby intersections and may, sometimes, even be detrimental to transit operations (for instance, for vehicles running ahead of schedule). Ideally, a really smart signal control system should exploit real time information about transit operations and general traffic conditions and adapt as efficiently as possible to changing conditions while minimizng disruptions in networks with coordinated signals. In the process of continually revising signal timings it should consider in real-time such factors as:

1. Traffic volumes at all approaches to intersections.
2. Queue lengths and potential spill-backs that might block lanes or intersections.
3. Expected arrival times of transit vehicles at signalized intersections.

4. Expected passenger occupancies of transit vehicles and other vehicles.
5. Deviations from schedule, i.e., how far ahead or behind schedule a transit vehicle is.
6. Deviations from proper service headways with respect to preceding and following transit vehicles.
7. Expected demand and wait times at downstream transit stations.
8. Expected arrival times of connecting transit vehicles at downstream transfer stations.
9. Expected delays to transit riders, motorists, pedestrians and other users, that are attributable to signal control decisions.
10. Expected vehicle operating costs attributable to signal control decisions.
11. Expected energy consumption and air quality impacts of signal control decisions.
12. Policy-based priorities that may be specified to create mode choice incentives, such as encouraging transit use.

An adaptive control system should also predict conditions well ahead of time and start the desired adjustments early rather than wait until the last moment, when options may be very limited. Still, it should be able to respond quickly to any new information, including major surprises. Thus, an ideal adaptive control system should have very good predictive capabilities as well as data collection, data processing, decision making and communication capabilities.

Unfortunately, existing signal control systems lack most of the characteristics and capabilities listed above. Although systems are beginning to emerge that can reasonably claim some adaptive control capabilities, they have very limited abilities to deal with real networks of coordinated signals and handle transit vehicles very primitively, if at all. There is a glaring need for adaptive control systems that can efficiently handle transit vehicles.

The need to control vehicle movements along transit routes arises because headways are naturally unstable. Given probabilistic variations in (a) dwell times at stations (due to variable number of boarding and exiting passengers) and/or (b) speeds along routes (due to variable traffic congestion levels), the natural tendency of transit vehicles is to bunch up in platoons. Thus, if a bus falls slightly behind schedule for any reasons, it will have more than the average number of passengers to pick up at the next station, which causes further delays and even larger abnormal loads downstream. Thus, it keeps failing further behind schedule. Conversely, the bus behind it encounters fewer passengers than usual and lower dwell times, allowing it to catch up with the preceding bus. Such bunching tends to increase as distances along routes increase, as scheduled headways decrease, and as demand and traffic variations increase. It is highly undesirable, since it can greatly increase wait times. (For example, if buses arrive at equal headways every 10 minutes, the average wait time is about 5 minutes. However, if buses bunch in four-vehicle platoons arriving every 40 minutes, the average wait time quadruples to about 20 minutes.) Transit operators devote considerable efforts and resources to preserving headways as uniform as possible and preventing or counteracting the natural tendency toward bunching. This has been relatively difficult control problem and a major motivation for the development of automated vehicle location (AVL) systems. A major part of the problem is that it is difficult to speed-up vehicles in congested traffic when they fall behind schedule. In such cases we may, reluctantly, also slow-up the following vehicles to preserve sufficient, and sufficiently uniform headways, or skip some stops. Signal pre-emption is a very promising way of speeding up transit vehicles.

It is very important that deviations from schedule (or from uniform headways) be detected and corrected as early as possible, by whatever means are available, before a problem gets out of hand. Hence, considerable sensitivity, reliability, precision and intelligence are desirable in the surveillance and control system.

Adaptive signal control can be used not only to stabilize headways and thereby minimize passenger wait times at stations. It can also be used, in conjunction with headway stabilization or without it, to help synchronize vehicle arrivals at transfer stations, and thus minimize transfer delays. In transit networks, transfers of passengers among routes may be used to (1) obviate the need for direct routes connecting all origin-destination pairs and (2) concentrate passengers on major routes to take advantage of high speed (and high cost) equipment. However, wait time at transfer stations seems particularly unattractive to users and may significantly deter transit usage. This wait time at transfer stations may be drastically reduced if the arrivals of vehicles from diierent routes can be synchronized or otherwise coordinated. Although successful timed transfers may greatly improve service in existing networks, major restructuring of transit operations may also be desirable to really take advantage of this concept's potential.

Two simulation models for bus dispatching control and for adaptive signal control have been developed and jointly tested in this study. The bus dispatching control model is developed for evaluating, and eventually implementing strategies along bus routes, Two major bus operation strategies, headway-based control and schedule-based control, are explored with the model. The headway-based strategy is to maintain proper bus headways in order to reduce bus bunching and passenger wait time. The schedule-based strategy controls buses toward keeping the original schedule instead of maintaining a desired headway. Buses are controlled to adhere to their own schedule regardless of how much bunching occurs.

The dispatching control model measures the system performance in terms of passenger in-vehicle time and wait time, bus travel time and headway regularity analyzes the effects of bus holding control and stop-skipping control on the effectiveness measures, while estimating the costs to users and suppliers and providing real-time information on bus movements and on-board passengers to control centers. Finally, this model can search for a combination of decision variables (i.e. simultaneously select the holding and skipping control values) that minimizes specified objectives.

The signal control model developed in this study is a pre-timed phase-based control procedure. The model evaluates a combined cost function of vehicle delay, number of stops, and on-board passenger delay, and develops a preset steady timing plan that minimizes the cost function.

With the two control models, several performance measures are analyzed at individual intersections and on bus routes. Holding and stop-skipping controls are analyzed and optimized based on a specified cost function. Headway-based and schedule-based control strategies are compared by various criteria of interest. Traffic operating costs and bus wait times are also evaluated through sensitivity analyses of parameters such as bus headways, bus delay costs, and signal timing.

The results from the dispatching control model also show that a headway-based strategy is preferable fpr minimizing wait time or headway deviations. Conversely, a schedule-based strategy is preferable for minimizing total cost, user cost, or user time. Analyses of signal control test results show that bus priority control is cost-effective for short-headway bus services but not beyond average headways of 40 minutes. Another finding is that longer than minimum feasible cycles may be preferable at short bus headways. Bus priority control with a minimum feasible cycle is found beneficial only for the longer bus headway services. A stability boundary is also suggested for identifying full bus priority conditions.

The integration of the two models is tested in a network. Traffic data and the simulated intermediate performance measures are communicated between the two models on a node-by-node basis. To compare the model with a no preemption condition, 250 sample buses on a two-way bus route are simulated. Numerical results show that the average bus delay time can be reduced by up ‘to 55% with bus priority control. With 5 minute bus headways, the combined operating cost for buses and other traffic is reduced by approximately 6% with priority control.

The concepts and models developed in this study can help transit operators to significantly improve the economic performance and service quality of transit services.

Table of Contents

List of Figures

List of Tables

Acknowledgment

The authors wish to thank Dr. Peter Su and Dr. Gang-Len Chang from the University of Maryland and Mr. Dennis Symes and Mr. Ron Fisher from the Federal Transit Administration for their advice and support in this work. This report was to a large extent funded by FTA University Research Project MD-26-7002.

Chapter 1 Introduction

1.1 Background

The quality of transit bus service concerns both operators and users. Bus routes may have many stops and signals which affect bus movements and operational efficiency. The need to control bus movements along bus routes arises because their headways are unstable. One natural tendency of buses traveling along a route is bunching up into platoon. This occurs because variations always exist in bus dwell times at stops and/or travel times along the routes. If a bus falls slightly behind schedule for any reason, it will have more than the average number of passengers to pick up at the next station, which causes further delays and even larger abnormal loads downstream. As a result, the bus keeps falling further behind schedule. Conversely, the following bus would have fewer passengers than usual and lower dwell times, allowing it to catch up with the preceding one.

There exist some measures of effectiveness (MOE's) to fit different planning and design goals. One of the main criteria in evaluating transit system, for both users and operators, is delay (or vehicle wait time) at signalized intersections. Accordingly, preferential treatment for transit vehicles is increasingly considered in transportation projects. Signal preemption has been considered for years as a feasible way to significantly reduce delay to transit vehicles. Some signal preemption operations, based on simple logic, have been implemented decades ago with much simpler technology than that available today. However, the simplified logic which was then feasible (e.g. immediate priority to buses in all cases) was not really adequate even in handling isolated, low-volume intersection interactions. For congested and sometimes even closely interrelated intersections, transit delays and related problems resulted from service unreliability become far more significant and intractable. In such circumstances, the usual simple control logic which automatically and immediately favors transit vehicles may cause severe problems to other traffic or environmental impacts. In addition to serious disruptions in a coordinated system of signals for nearby intersections, these controlled manipulations may also be detrimental to transit operations. Therefore, the issue of tradeoffs between the social benefits and costs from signal controls for bus priority has become a major one in traffic management.

Control options may improve the reliability of transit bus service and reduce their riders' wait time at the cost of increasing passenger in-vehicle time and bus travel time. With such limitations, a comprehensive analysis of transit vehicle operating performance is needed. Ideally, several control options could be coordinated to maintain the regularity of bus movements on routes. The options could be those treatments only good for bus operations either at some nodes (e.g. bus stops or signals) or on some links (e.g. bus exclusive roads or contra flow lanes). In fact, an ideal adaptive control system could not only exploit real time information about transit operations and general traffic conditions but also adapt itself as efficiently as possible to changing traffic conditions in the networks. Therefore, it would be able to: (1) Predict conditions well in advance and activate the desired adjustments as early as possible, rather than urgent by the last moment when options may be very limited. (2) Quickly respond to any new information, including major surprises, in order to be fully adaptive between traffic conditions and controlled actions. (3) Possess strong capabilities in data processing, decision making, signal transmission and execution. Given these characteristics, the designed control mechanism could function itself to meet the management goals as well as balance the tradeoffs between public interests and private preferences.

Unfortunately, existing signal control systems are usually deficient in most of the above features and attributes. They still have limitations for dealing with varied problems, such as signal coordination and efficiency of transit operations, in diierent transportation facilities. Based on these points, the development of an adaptive bus control model is proposed in order to effectively operate transit vehicles. The results show the success of preemption control in reducing time cost and operation cost as well.

1.2 Objectives and Scope

Bus movements are usually mixed with other traffic. Inevitably, unexpected disturbances may be brought about by internal and external factors of the traffic flow. The external factors may involve elements such as characteristics of transit buses, roadway or traffic flow impedance, random delay due to signals, and unusual passenger demands. The internal factors may include elements such as organization and management, scheduling, and bus assignment. Combinations of such elements might cause those bus movements to be more complex and uncontrollable. Therefore, finding treatments for ameliorating bus operations in mixed traffic is the major goal of this study. Two major objectives are proposed to frame the our intention. One is to explore some efficiency problems that transit buses might face when they operate through signals and stops. The other is to develop an integrated control approach for transit vehicles moving along their route.

As emphasized previously, there are many elements affecting the operational performance of transit vehicles. These elements may come from either links or nodes. The influence of those link elements could make traffic behavior, such as lane changing, and acceleration / deceleration maneuvers, too complex to be investigated. Thus, bus operations are treated only at nodes in this study. Two major node controls, bus dispatching at terminals / stops and signal preemption at intersections are regarded as the most effective ways for reducing cost of both transit vehicles and their passengers. The functions of these two control strategies constitute the critical issues in the following sections. Moreover, their joint applications are also simulated and explored in this study.

1.3 Research Methodology

To select and test the proposed approaches of node control, two preliminary tasks are carried out. Firstly, some related studies and existing field control models are briefly classified and reviewed. These cover treatments of bus operation at stops, bus priority at signals, and existing adaptive signal control models. Secondly, a bus route with 6 signals and 10 bus stops is proposed for a case study.

Treatments of node control at bus stops and signals are modeled individually, In the development of control models for bus stops, impacts that could delay buses through intersections, such as signal blockage and turning movement disturbance, are omitted. Instead, log-normally distributed travel times are generated to simulate bus movements between nodes. For bus priority, models with pre-specified signal timing options are developed for isolated intersection control. Optimization procedures are executed when any information regarding bus movements and/or locations is received. Some characteristics of the models for bus stops and signals are analyzed separately before they are combined and jointly implemented.

The joint application would especially emphasize the linkage of the above node control models. Bus dispatching time from any stop or discharging time from any signal has direct effects on control decisions at all downstream nodes. Thus, the control decisions for both stops and signals are mutually affected when buses travel along the route. Through the model simulation and sensitivity tests, a few results can be obtained by controlling some critical variables such as bus headways, service types (emergency levels), and signal timings. With these findings, a field control plan is proposed for further tests and applications.

The report contains seven sections which are briefly described below:

1. Introduction

This briefly describes some efficiency and delay problems of bus operations that might occur at either bus stops or signals. It also defines the subject and its scope of this study.

2. State-of-the-art review

Bus control models and methods at either stops or signals are explored and discussed. In addition, some contemporary adaptive signal control models and their concepts are also reviewed.

3. Control models for bus movements at bus stops

Two major control strategies, headway-based and schedule-based controls, are proposed for bus operations at stops. With each control strategy, two options of bus dispatching, holding or skipping control, are tested separately in the simulation runs. Cost-based statistics extracted from those simulation runs are classified and compared in pairs to make final control decision.

4. Control models for bus priority at signals

A simple control logic is proposed to immediately adjust current signal phasing for treating the requests of bus passage. Traffic operating cost including passenger car delay, total vehicular stops, and expected bus wait times are calculated and minimized to find out the best phasing adjustment. Some service measures such as headways, bus wait cost, and signal timings are also used to test the possible effects the control model may cause.

5. Controls for transit vehicles along signalized bus routes

Bus movements along a signalized route are simulated by combining the previously developed node control models, i.e. bus dispatching control model and bus priority model. Test results obtained via such simulation runs are systematically analyzed and cited to support the conclusions.

6. Adaptive control test plan

Some limitations of an adaptive transit system and their measures of effectiveness under study are specified. A series of test plans for signal control, travel speed control, and bus dispatching control at transit stations, are developed for testing the models proposed in the study.

7. Conclusions

The section states main findings of analyses and suggests future work concerning some incurred problems of the research.

The entire research procedure is shown in figure l-l.

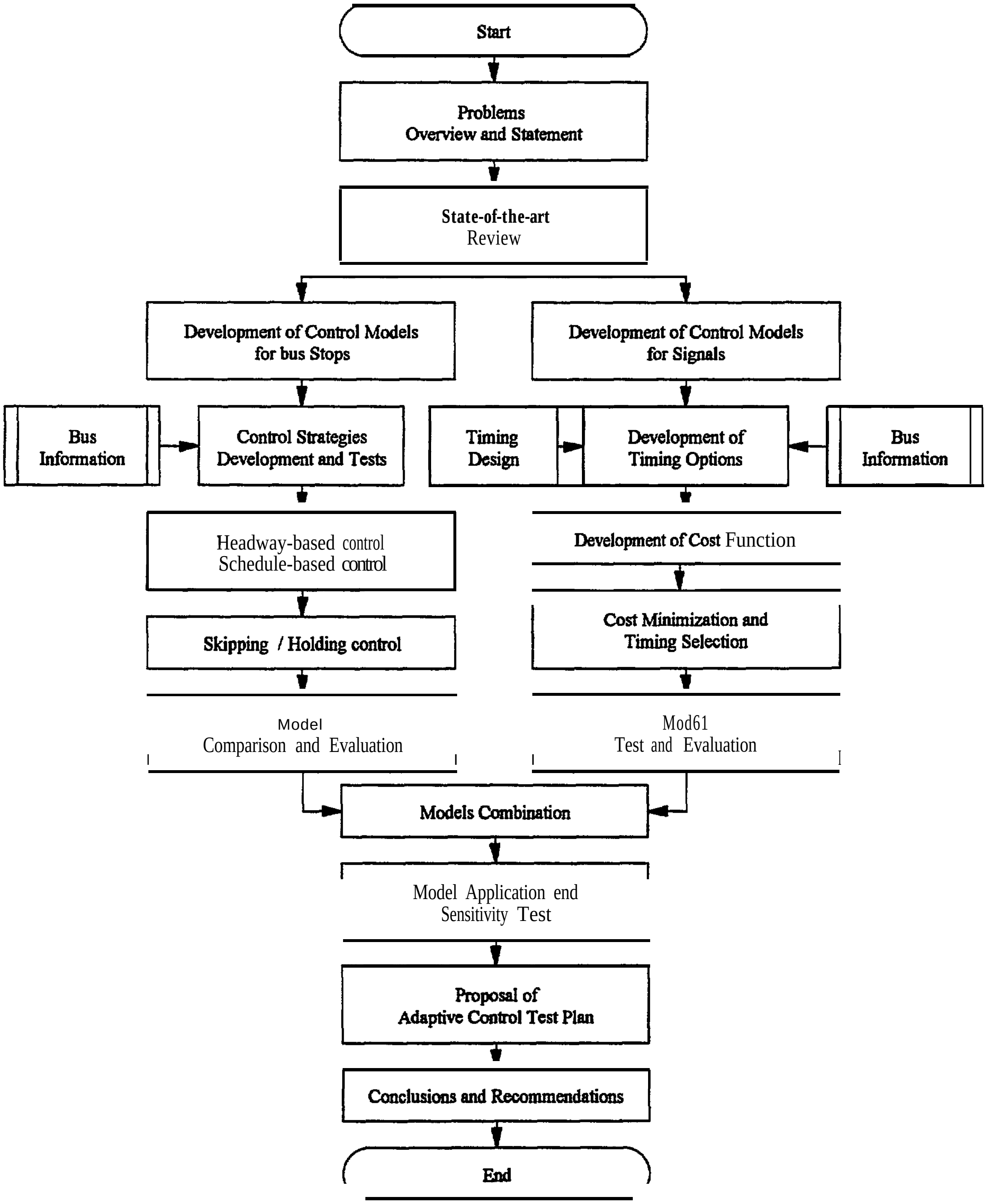

Figure 1 - 1 Flow chart of research procedure

Chapter 2 State-of-the-art Review

2.1 Route Controls for Bus Operations

The major reliability problems in transit service are the platooning of vehicles along routes and poor connections at transfer points. The platooning and poor transfer connections can be traced to excessive variability in either link travel times between stops or dwell times at stops. Therefore potential control strategies should be focused on reducing one or both of these sources of variability. In a broad sense, the major objectives of control strategies are to keep platoons from forming or to break them up after they have formed, and to ensure proper arrival times at transfer points.

Accordingly, four classes of strategies can be used for improving the reliability of bus transit service. They are: (1) reduction of the number of stops, (2) signal preemption, (3) provision of exclusive right of way, and (4) vehicle dispatching controls.

2.1.1 Reduction in Number of Bus Stops

Long routes are more likely to develop schedule deviation problems. As a bus travels farther from its origin its deviation from schedule tends to increase along the route, since more and more stochastic factors are cumulated. Shorter lines are usually easier to keep on time because opportunities for recovery (i.e., layovers) occur more frequently. For maintaining the reliability of uncontrolled transit service, routes should generally be as short as possible. Reducing number of stops (i.e., increasing stop spacing) or breaking long routes into segments can alleviate service reliability problems [l, 2, 3, 4, 5, 63. However, shorter bus routes and longer stop spacings may increase passenger transfer times and access times, and frequent layovers of buses at terminal stations may increase bus wait time.

2.1.2 Provision of Right of Way for Buses

It has been found that reserved lanes for buses improve the operation of buses, especially during peak periods. There is considerable empirical evidence from the United States and other countries that reserved lanes can improve both average transit speeds and reliability [3, 7, 81. Nevertheless, bus lanes result in lane blockage at bus stops and formation of weaving areas near intersections. These two disadvantages reduce the benefit obtained from bus lanes. In general, the main factors to be considered when deciding whether a bus lane is desirable include the street width (the number of lanes), the volume of buses in the peak period, the degree of congestion and the bus occupancy.

2.1.3 Control of Bus Movements along Routes

2.1.3.1 Bus Schedule Adjustment

Schedule adjustment might include implementing tighter or looser schedules. To achieve reliable bus service, realistic schedules are required. Schedules that are too fast will result in poor schedule adherence, while schedules that are too slow will result in long travel times for both passengers and buses. Bus dwell time and bus running time are two of major factors in setting bus schedule. A realistic schedule requires empirical data on actual travel times and dwell time and considers the stochastic characteristics.

Abkowitz and Engelstein (1982) examined transit running times at various times of the day, in diierent directions of travel, and at different points along route with empirical data from Cincinnati, Ohio. They found that transit running times are highest and most variable during the afternoon peak period. Regardless of time period, it is apparent that variation in running times increases with distance from route origin so that service deteriorates as the vehicle proceeds downstream [9].

An analysis of bus travel times and speeds was conducted in a cross section of U.S. cities by Levinson in 1983 [l0]. Three basic analyses were conducted: (a) bus and car speeds were compared; (b) bus travel times and delays were estimated from various field studies; and (c) bus travel times were derived based on dwell time, traffic congestion, actual acceleration and deceleration rates, and distance between stops.

Abkowitz and Engelstein (1984) developed a linear regression model to estimate the mean running time. In the model, link length, passengers boarding, passengers alighting, percentage off-street parking signalized intersection, time of day, and direction of travel were chosen as independent variables. Their analysis of bus running times uses bus operations data collected from Queen City, Cincinnati. This model showed that mean running time is highly influenced by trip distance, boarding and alighting, and signalized intersections and to a lesser degree by parking restrictions along the route, time of day, and direction of travel [11].

Guenthner and Sinha (1983) developed a mathematical model for estimating bus dwelling time at bus stop using data from Milwaukee and Lafayette. In this suggested model two basic assumptions were made. They are: the number of passenger boarding and alighting at each bus stop follows a Poisson distribution, and the passenger demand is uniformly distributed along a bus route. Bus dwelling time consists of the delay time for the stopping and starting maneuver of a bus, the delay time for those stops with 24 or more boarding and alighting, and the dwell time for stops with 23 or less boarding and alighting. Using this model, authors revealed two major findings: (1) An increase in posted stops along a low-demand route will have only a minor effect on bus operating speed and reduce the user's walking distance; (2) Additional posted stops along a high-demand route will save walking distance at the cost of greater m-vehicle travel time [12].

2.1.3.2 Adjusting Service to Desired Headway

Such controls may include holding early buses and skipping stops to adhere to schedule or maintain more equal intervals between successive buses. The previous studies mainly focused on determining threshold value of holding and stop-skipping controls, and identifying optimal control points.

Osuna and Newell (1972) used an analytic method to determine the optimal holding strategy at a bus terminal. In this model, they adopted a uniform distribution of passenger arrivals and assumed that bus bunching did not occur, and that buses had sufficient load capacity. Their analysis focused on holding control strategies for a simple bus system with one or two buses. This optimal holding threshold value a* is suggested as follows:

For a system with a single bus:

$$a^* = E(w)\big|_{a=a^*} \qquad (2.1)$$

For a system with two buses:

$$a^* = \frac{1}{2}\left\{E(T) - \left[\frac{3}{2}Var(T)E(T)\right]^{\frac{1}{3}}\right\} \qquad (2.2)$$

where

$E(T)$ = average round trip travel time of bus

$E(w)$ = average wait time of passengers

a^* = optimal holding threshold value

$Var(T)$ = variance of bus travel time

They suggested that one should not apply control in anticipation of bad situations, but wait until they happen [133.

Koffman (1978) simulated the movement of a single-direction bus route. In his model, headway-based control was implemented. Two holding threshold values and two skipping threshold values are compared. Bus-priority signal control and dispatching uncertainty are considered. Whether bus preemption was implemented only depended on the bus on-time performance. Koffman concluded that reducing the dispatching uncertainty and using signal priority control could significantly improve wait time [14]. However, no interpretation of the holding and skipping control results was provided.

Turnquist and Blume (1980) analyzed the holding strategies with a probability model of vehicle arrival time at the control stop. They concluded that for control to be effective, the optimal minimum headway after control must be greater than the short headway before control, or it does not pay to control at all [IS]. In the analysis, the wait time of passengers at a specified bus stop was taken as the main criterion. They indicated that the major costs of such a policy were borne by passengers who were already on the vehicle, since they were delayed when the bus was held up. Thus, the implementation of holding control strategy makes some passengers better off at the expense of others.

Turnquist (1978) concluded that a schedule-based control could be particularly useful on suburban routes or in other instances in which headways were quite large. The effectiveness of headway-based controls depends on identification of an appropriate control point along the route. He suggested that the control point should be located as early along the bus route as possible [3].

Abkowitz and Engelstein (1984) developed an analytic model to determine the optimal control points and threshold values according to the following cost function:

$$TW = \sum_{i=1}^{j-1}(n_i \overline{w}_i) + [b_j d_j(x_0)] + \sum_{i=i}^{N}(n_i \overline{w}_i) \quad \text{(2.3)}$$

where

TW = expected total wait time on route,

j = the control stops,

x_0 = threshold,

n_i = number of passengers boarding at stop i,

b_j = number of passengers on board at stop j,

w_i = average wait time at stop i,

N = total number of stops on route, and

$d_j(x_0)$ = expected delay at the control stop for the threshold of x_0.

They found that the location of the control stop is quite sensitive to the distribution of passengers boarding at stops. Generally, the control point occurs just before a group of stops at which many passengers are boarding. Thus, more passengers enjoy a reduction in the wait time, because the headway variation is mainly reduced at stops that are close to the control point. If the number on board is small, it is more likely that the threshold value will be large. The threshold value and the location of the control point are interrelated and they are dependent on all the input parameters in the algorithm [11].

Abkowitz and Engelestein (1986) concluded that headway-based control was suitable for routes operating with short and uniform headways. When headways are short and uniform, it is assumed that passengers arrive more randomly at stops and that they are mainly concerned with the headway rather than the schedule. Operators are concerned about keeping vehicles evenly spaced so that vehicle availability remains stable. Abkowitz and Engelestein also considered that schedule-based control was suitable to routes that have long and/or uneven headways [16].

Abkowitz and Tozzi (1986) developed a mathematical model to investigate the impact of five boardmg and alighting profiles on the effectiveness of headway-based control. These profiles specified that passengers boarded at the beginning and alighted at the end of route, boarded at the beginning and alighted in the middle and the end of route, boarded at the beginning and alighted in the middle of route, boarded and alighted uniformly along route, boarded in the middle and alighted at the end of route, respectively. They suggested that implementing headway-based control for uniform boarding profiles may be more feasible on routes with heavy ridership [17].

Seneviratne (1989) developed a simulation model to examine the performance over time. In this model, the control points on the route are optimized according to a specified criterion: the maximum permitted 60 seconds of headway standard deviation. At the end of each simulation set, the point on the route where headway standard deviation exceeds 60 seconds is identified and a control point is placed at the preceding stop [18, 19].

Abkowitz and Lepofsky (1990) presented the results of implementing real-time headway-based reliability control along candidate bus route operated by the Massachusetts Bay Transportation Authority in 1987. This procedure involves holding buses at a control point on the route until a prescribed minimum headway is achieved. They concluded that headway-based control strategies were applicable to high-frequency transit routes where headways are sufficiently short so that travelers arrive randomly at bus stops without consulting a schedule. In their model, the performance measures of interest related primarily to passenger waiting and delay times (on board delay), vehicle running times and headways [20].

2.1.3.3 Adding Reserved Vehicles

Bus service reliability problems worsen as buses proceed along a route. If dispatching at the route origin is on time, the headways will be reasonably regular at the early spots along the route. To dispatch at the origin on time, it is very effective to give drivers more recovery time (layover) to ensure the bus leaves on schedule, or to use spare buses when the scheduled bus can not get to the key points or when excessive crowding occurs. Sometimes a run is delayed because of a defective bus, an inexperienced operator, or other unusual circumstances. When this occurs, it may be desirable to add a reserve bus to fill the gap at that point in order to achieve a better distribution of passengers on each vehicle. Houston, Seattle, and Toronto use reserve buses at key points along the route [21]. The tradeoff should be made between improvement in regularity and increase in bus supplier cost due to additional buses.

Over the past several decades, considerable progress has been made toward a better understanding of transit reliability control. However, previous studies have had the following weaknesses:

(1) Ignoring the dependent relation between bus arrival and control strategies when a bus arrival time distribution is used;

(2) Assuming an infinite passenger capacity in each bus. This assumption neglects the influence of bus capacity (bus frequency and seat capacity) on regularity of bus movements.

(3) Implementing controls at selected control points. Such control has two major weaknesses. First, only part of the cost can be considered. Second, such control is less effective. If a bus is very late when it arrives at the selected control point, it would be diicult for the bus to return to its expected trajectory even if it is instructed to skip stops.

(4) Lacking comprehensive analysis of the effects of control strategies on passenger wait time, passenger in-vehicle time, bus travel time, user cost, and supplier cost.

(5) Lacking comprehensive comparisons between headway-based control and schedule-based control, as well as optimization for combinations of holding control and skipping control.

2.2 Preemption for Transit Vehicles at Signals

Around 1975, experiments were conducted in the U.S. and several European countries to test various methods of minimizing bus delays at intersections [22]. By taking advantage of the IVHS technologies, most transit experts have widely conceived the applications of transit bus preemption as a major tool in alleviating urban traffic congestion problems.

2.2.1 Related Real-time Models

Effectiveness of transit operations at signals is an important factor in modeling control system for road networks. The control could improve the productivity of an intersection by increasing its throughput or by decreasing total person-time of delay or related vehicle operating costs. Theoretically, a real-time signal control model should be processed more readily in accommodating with the non cyclical effects of bus operations than a fixed-time control model does. However, in fact, both real-time and fixed-time signal control models fail to treat on-line transit operations effectively. This is mainly due to the difficulties of collecting and processing on-line data concurrently. With the strengths and deficiencies of treatments for contemporary transit operations, some analytical models and simulation models for bus priority control at signals are briefly reviewed and discussed below.

SCATS, the "Sydney Coordinated Adaptive Traffic System", developed in Australia as a real-time adaptive control model, has been widely tested for promoting traffic operation efficiency [23, 24, 25, 26]. In addition, a project entitled SCRAM (Signal Coordination of Regional Area in Melbourne) has been proposed to enhance SCATS in facilitating public transport priority [27]. With two categories of priority provision (passive priority and active priority), accompanied with appropriate timing strategies, SCRAM really extends the signal control functions for public uses. For the level of passive priority, historical data on transit vehicle (tram) operations and behavior are used to predict requirements for tram priority. This level focuses on reducing the major sources of operating delay (approach delay and stop delay) to the transit vehicles. Thus, several timing design features such as using minimum cycle length, providing special phase design for exclusive tram movements, and providing green phase progression band are specified in this level. The main function of the active priority is to selectively detect the transit vehicles (trams) in the traffic stream and directly adjust signal timings for them. To offer higher quality of bus service, several strategies such as executing either green extension or phase early cut-off, providing special phase design within multiple phases to smooth tram operation, and suppressing non-tram phase to quickly serve a tram phase, are considered in this level. However, although the system does provide reasonable control features, it still fails to treat two or more transit vehicles coming from different approaches at the same time.

SPPORT, the "Signal Priority Procedure for Optimization in Real Time" model is primarily developed to incorporate methods of traffic responsive signal control and operational control of transit vehicles in traffic flow [28,29]. By considering period-based events, the traffic-responsive tool utilizes a fairly reasonable approximation to treat each signal with large amount of uncoordinated traffic demands on all competing approaches. SPPORT is said to be real-time for it could continually detect and use traffic information to update current signal plan every five seconds. To estimate departure times from intersection, the model employs a FIFO queuing discipline with vehicle headway under saturated condition as its service time. SPPORT can order important events by priority in order to allocate green times. The higher an event is on the list, the more likely it is to receive a green phase. The program is able to pre-evaluate each of the phase sequences generated from the respective priority lists by using pre-defined cost function. Also, it can dynamically select the most promising plan on-line for immediate short-term application. Although some features are especially required to develop the model, both control concept and timing strategy are most critical to make the model's efficiency.

For the control concept, SPPORT is capable of (1) Considering events for a period of time rather than the events at a specific point of time. (2) Constructing up to 11 priority levels (such as queue length, queue served condition, load / unload approaching streetcar, load / unload approaching transit buses, emergency vehicles, and maximum green time) to further develop time-weighted priorities for evaluation. (3) Applying the B/C ratio to compare the time-weighted priorities calculated from collected traffic data on different intersection approaches. For the timing strategy, (1) if an event has priority P at time T_1 and lasts until time T_2, the time-weighted priority for this event will be $(T_2 - T_1)P$. Based on the priority, a decision index TWP can be calculated as

$$TWP = \sum_{i=1}^{N_g}(TWP_g) - \sum_{i=1}^{N_r}(TWP_r) \quad \text{...}(2.4)$$

where

N_g = Number of green approaches

N_r = Number of red approaches

TWP_g = Time-weighted priority for a green approach

TWP_r = Time-weighted priority for a red approach

(2) If the TWP value is equal to or greater than 1 .O, the signal phase will not be switched. Otherwise, it will be switched. These considerations characterize the entire model's

operation. However, there also exist some problems to the model such as limited capability of foreseeing uncertain traffic patterns dynamically, high cost of installing high-speed computers as well as their communication system, and long computation time in optimizing solutions.

In addition, a system called UTOPIA, the "Urban Traffic Optimization by Integrated Automation", was initially designed by the Fiat Research Center in Italy and has been tested in Toronto and Turin. It mainly considers control of private vehicles together with a comprehensive public transport operations within a large scale, hierarchical decentralized traffic adaptive control system. Problems are classified into two levels, Intersection level (lower level) and Area level (decision level). The area level traffic model predicts O-D for passenger cars based on historical data and real-time information collected from local intersections. Then, a cost function considering delay to intersection traffic flow, public transit buses, and the whole area decision policy is optimized at local level [30, 52].

For the intersection level, UTOPIA could: (1) Utiliie its microscopic model to simulate traffic flow at a signal. (2) Determine the signal setting to get some traffic performance index such as delay time to passenger cars and transit vehicles, vehicle stops, queue length, and deviation from signal setting decided in the previous iteration. For the area level, the model can: (1) Analyze area-wide traffic data and make predictions for main street flows in time. (2) Apply its internal macroscopic model to entire area network and traffic counts. (3) Optimize the total travel time with constraints of average speed and saturation flows. Yet, practical applications of the model have shown that the use of average link travel time from upstream detectors may directly impact system prediction validity and optimization performance. Also, the reliability of O-D prediction is still low for practical uses.

2.2.2 Enhanced Off-line Control Model

The development of the “PREEMPT” computer simulation model is intended to present and test the feasibility of using bus preemption as a tool for transit operation management [3 1]. The program was initially operated along an urban arterial to reduce travel time and to improve overall travel speed without any on-board quick-response equipment. With a built-in elasticity-based demand algorithm, accompanied with three entities (① Fleet size, headway, and cycle time, ② operating cost and revenue, and ③ elasticity-based demand function), the model could explore the possible effects of improved quality of services and fare changes on overall traffic operating cost.

PREEMPT uses “need” and “eligibility” criteria (about 5-10 seconds) to qualify a bus preemption decision, In addition, it also proposes three strategies to arrange the signal timing: (1) green extension, (2) red truncation, and (3) red interruption. The desired result is fully determined by linking three entities:

(1) Fleet size, headway and cycle time

-Cycle time $C=T_d+T_s+T_c$ ....(2.5)

• Number of bus required and fleet size:

$N_v \geq (D_p \times C)/(V_C \times 60)$ .... (2.6)

-Headway $H=C/N_V$ ....(2.7)

where

D_p = Hourly passenger demand

V_C = Bus capacity

T_d = Driving time

T_s = Board / unboard time

T_c = Layover time

(2) Operating cost and revenue:

$D_p = K(T)^{Y1}(P)^{Y2}$ ....(2.8)

where

K = Coefficient

T = Travel time

P = Cost of travel

Y1 and Y2 = Time elasticity of demand and cost elasticity of demand

(3) Elasticity based demand function

$$FAC = \$1.025\,(X) + \$21.03\,(Y) + \$80516\,(Z) \quad \text{.................................} (2.9)$$

where

FAC = Fully Allocated Cost

X = Annual total vehicle miles

Y = Annual total vehicle hours

Z = Number of buses required to provide peak service

Unfortunately, there are still some limitations in the simulation model such as no capability of determining the economical fleet size and lack of a model validation function throughout the actual deployment of preemption hardware.

2.2.3 Other Simulation and Delay Models

In reviewing the TRANSYT program' the entire traffic system is categorized into two dimensions called "BUS TRANSYT" and "BASIC TRANSYT" [32, 46]. This program was applied to obtain appropriate signal offsets and phase splits at a hypothetical site with diierent levels of intersection volumes. The tests of model's performance measures have shown that a bus-actuated control system especially suits low bus flow conditions while a fixed-time control gives a better performance measure with high bus flow conditions. In addition, several bus control strategies were regulated in this program to provide deferent levels of priority to transit buses. The authors conclude that a bus-actuated control system operating during the major road green stage would show an improvement in the calculated P.I. over a fixed-time system with offsets and splits given by "BUS TRANSYT". It also shows that the performance indices produced by a bus-actuated system are higher than a fixed-time system if considering the compensation to the traffic on the side street.

With regard to the trade-off analyses of road user costs, some network-wide models have been run to evaluate travel time delay and fuel consumption of all vehicles. UTCS/BPS (Urban Traffic Controls System / Bus Priority System) and NETSIM (Network Flow Simulation for Urban Traffic Controls System) computer simulation models were utilized to estimate the measure of effectiveness, i.e., travel time delay, for both preemption and non-preemption cases [33]. Through instantaneous data generation, the fuel consumption rates and emission rates can be calculated for incremental benefit-cost analysis. Even though the results indicate a bus preemption system would be cost effective on a network basis, the effectiveness varies among different signals in the network.

The methodology of analytical delay models under bus signal preemption was also proposed for application previously. One possible strategy is the priority treatment, which is also aimed at improving the capacity of intersections, of buses at signalized junctions. Jacobson and Sheffi have developed delay models for testing traffic impact with signal bus preemption in 1980 [34]. Their test results showed that bus priority could greatly reduce hourly person-seconds delay of operation at high bus occupancy and high flow rate conditions. Moreover, from the model analyses, they suggested that the benefits of bus preemption be increased by properly adjusting several design parameters such as signal cycle and preemption phase duration as well as some non-preempted parameters. Based on their model, quite a few directions of signal preemption studies were recommended such as the correction for queue length-dependent service rates for passenger car flows, multiple preemption control at a given cycle, and modeling the arrival process of vehicles in batches.

Generally, when performing signal control analyses, both the design of a bus priority scheme and such a mechanism be expected to show play a very important role in many studies. Most of analytical methods or simulation models developed for decades usually use a common measure of delay incurred while traffic passes through signals with either uniform arrival rates or complex variable flow patterns [35, 36]. For accurately estimating the traffic measure, it is necessary to conduct more field investigations and test the signal adjustment plans. By means of further benefit-cost analysis, the purpose of signal adjustment could be satisfied for both private and public vehicle controls.

2.3 Contemporary Adaptive Traffic Signal Control Systems

The current signal control hierarchy can be briefly divided into three types: (1) centralized control system' (2) two-level distributed control system' and (3) multi-level distributed control system. A centralized control system connects all of its local controllers to a central control unit. The unit only performs switching tasks and simple data processing. A two-level distributed control system makes all of its local controllers execute intelligent control instructions which are sent by the central computer. A multi-level distributed control system contains three levels: local controllers, regional computers, and central computer. The first two levels mainly perform traffic responsive control task while the third level executes command and monitor controls.

Based on the above hierarchy, together with signal timing strategies and control period, the current signal systems are categorized into three. They are (1) short-term network control systems, (2) cyclic network control systems, and (3) acyclic network control systems [37].

2.3.1 Short-term Network Control Systems

The UTCS family developed by FHWA in the early 1970's is the most famous one of this system. It uses off-line or on-line optimization models to respond to the traffic variations [38]. The UTCS has been developed up to 3rd generation. However, the evaluation of UTCS showed that the 1st generation one can outperform the others. Such conclusions have motivated the development of the 1.5 generation system in the early 1980's. The 1st generation system can pre-store an off-line timing plan by using its optimization control model. However, in the 1.5 generation system, the optimal timing plan can be generated by a specific feature selected by the users. In the 2nd generation system, the optimization procedure is developed based on the control module in SIGOP II to save computation time. The 3rd generation system has the capability of varying its background cycle length generated by SIGOP II within a short period 3-5 minutes. Also, its queue management control model can provide area-wide signal timings for saturated condition.

The CALIPE system is developed to improve the limitation of classical central control strategies so as to perform an on-line computation and update the adaptive timing plan on a 6 minutes basis. CALIFE uses the structure of 2nd generation UTCS to operates its traffic model and applies a model derived from TRANSYT-7 to perform its system optimization procedure. The main advantages of CALIFE are its capability of upgrading a classical signal system to conduct an on-line control functions, considering the possible disturbances in the transition process. However, the system only considers stable turning movements and saturation flows over a long period in the reconstitution step [39,40].

The CLAIRE is an European prototype system for treating congested traffic conditions. The expert system utilizes the UTCS as an I/O tool to make control decisions or modify control actions. It features on-line monitoring of traffic congestion and off-line learning control actions. A significant benefit of the system is its abiity to assimilate and process more information than 15-20 CCTV scanned by one operator.

2.3.2 Cyclic Network Control Systems

The major function of a cyclic network control system is to adjust signal operations based on the cycle-by-cycle flow variation. Signal cycle, green duration, and offset are optimized based on the performance measures such as stops and delay calculated by the internal traffic models. Two well-known systems, SCOOT and SCATS, are reviewed in this section.

The development of SCOOT, Split, Cycle, Offset, Optimization Technique, was firstly completed by British TRRL in the late 1975 and tested in the early 1981. It is developed to achieve full responsive control either at an isolated signal or in a signalized network. The central control system performs its adaptive control depending mainly on the interaction between a central computer and a local controller. It uses a TRANSYT like concept to predict real-time signal control effects and provides short- or long-term traffic information for system management purposes.

In SCOOT, an optimal timing plan evolves gradually by optimizing three parameters: green splits, cycle length, and offset. The whole optimization concept is developed with three main considerations: (1) no sudden change in timing transition' (2) no mechanism to predict traffic in several minutes, and (3) no sensitivity of detector false effects. Together with the built-in timing optimisers, the SCOOT feature has been proven capable of adapting medium traffic congestion conditions. However, the centralized system requires many computers and detectors to monitor the entire area under control. Also, the null intelligent local controllers may directly affect SCOOT to develop a local microscopic adaptive control strategy [41, 42, 43, 44, 45, 46].

The SCATS, Sydney Coordinated Adaptive Traffic System, was commenced by the Australian Road and Traffic Authority in the early 1970's. The system employs a central computer, regional computer, and local intelligent controllers to perform a large-scale network control. The regional computer can execute adaptive control strategies without any aid from the central computer which only monitors the system performance and equipment status.

SCATS employs a strategic optimization algorithm and a tactical control technique to complete the whole system optimization issues. The optimization philosophy contains four major modules: (1) cycle length optimiser, (2) split optimiser, (3) internal offset optimiser, and (4) linking offset optimiser. The system features and the intelligent design concept enable SCATS to expand easily and suitably for controlling any size of traffic area. However, the system requires a significant involvement from the traffic engineers to define the strategic detectors, the split plan' the offset plan, the cycle time, and the voting criterion [23, 24, 25, 26].

2.3.3 Acyclic Network Control Systems

An acyclic control system attempts to provide its optimal decision in l-5 seconds. Consequently, the control is not to determine the optimal cycle, splits, or offset, but to solve the best control sequence for adapting a very short-term demand variation. Such an algorithm requires an extensively simple computation process to quickly reach the solution. Some well-known approaches, such as Miller's algorithm (1963), Bang's Traffic Optimization Logic (TOL, 1976), MOVA (1988), OPAC (1980), and SAST (1988), have been developed to control isolated signals. Furthermore, a few network control features were also developed in the mean time. Two of them are reviewed in this section [47, 48, 49].

Research for developing PRODYN was initiated in 1978 and has been tested in both isolated intersections and networks in ZELT (Zone Experimentale et Laboratories de Traffic de Toulouse). The promise of this system is attributable to its sophisticated traffic state model and its dynamic on-line optimization techniques [50, 5 1].

In the demand prediction model, vehicle arrival times at the stop line are predicted for the next sixteen 5-second time intervals (time steps). This would cover a 75 seconds time horizon. An internal queue model assisted by those upstream detectors, can estimate queue length based on vertical queues, arrivals and discharge rates.

In the system optimization model, only the first step information of those siieen time intervals is implemented as an adaptive control input. The optimization criterion is to minimize the sum of delays over the time horizon. At the intersection level, the optimization model is to minimize delay by using improved forward dynamic programming with the constrains on maximum and minimum greens. At the network level, the network coordination optimization is performed by a decentralized control structure. The procedure includes: (1) simulating a specific intersection output for each time step as soon as the intersection controller finishes its optimization over the time horizon, (2) sending the simulation output to each downstream intersection controller, and (3) each of the downstream controllers uses the output message at the next time step to forecast arrivals.

CRONOS (ContROl of Networks by Optimization of Switchovers) was proposed in 1992. It incorporates several functions, such as modeling the over-saturated traffic conditions, monitoring traffic conditions at signals and on links, providing efficient optimization approach, in the real-time algorithm. The proposed optimization approach, with its polynomially increasing n^2 complexity, has shown a capability to perform real-time computation. Two major parts of CRONOS are described below.

Firstly, the traffic prediction model can take into account the queue spatial extension in each control link based on real-time image base detection and its past information. In addition, it can re-actuate and memorize the left-turn vehicles stored in the intersection at each time step in order to model the departures from the links. Secondly, CRONOS applies a rolling-time horizon (80 seconds) concept and a revised Box algorithm in the system optimization process. The optimization criterion is to minimize the queue delay and travel time between signals. The improved Box algorithm and optimization process give the model several advantages: (1) obtaining a near global minimum solution, (2) considering both traffic spill over and complex left-turn maneuvers, and (3) providmg only n^2 computation complexity.

According to the test results, CRONOS can perform coordination of three successive intersections. In addition, due to the successful use of image base detection techniques, the system has the capability of predicting real-time horizontal queue evolution and congestion level at an intersection. However, the high implementation cost might become the main obstacle for its application in a large-scale traffic network.

Chapter 3 Control Models for Bus Movements at Bus Stops

The purpose of this chapter is to develop a control model for bus movements along a bus route. With such a model, bus movements are simulated and bus control strategies are analyzed and compared. This chapter presents the description of the simulation model, analysis of experimental results, comparison of control strategies, and conclusions.

3-l Description of Bus Movement Simulation Models

A computer simulation model has been constructed to investigate the characteristics of bus movements and the relative merits of several control strategies. In the model, a bus route traversing an assumed urban area is modeled. The bus route consists of 10 bus stops (including origin and destination terminals) and 6 signalized intersections. The discussion will concentrate on the peak-hour case (baseline case) during which there are higher bus frequencies and higher load factors. Low load and high frequency cases, as well as low load and lower frequency cases are also considered in the sensitivity analysis section.

The computer simulation model consists of several subroutines, which are combined in two streams: passenger stream and bus stream (figures 3-1 and 3-2).

3-I-1 Description of Model Input / Output

For the model input, the following conditions are presumed:

1. Passenger arrival and departure rates along a bus route during the period of interest have been obtained. During this period, the passenger demand is approximately stable (figure 3-3).
2. Bus service frequency is pre-planned by the bus operator. In the base-line case (peak period), the specified bus frequency is 12 per hour. Suppose that the maximum load capacity of available buses is 80 passengers (seat capacity 53, and standing space 27) [53]. Thus, this bus system can serve passengers at a load factor of 0.76 eastbound

and 0.51 westbound at the critical (maximum) occupancy point. A plot of occupancy along the bus route is displayed in figure 3-4.

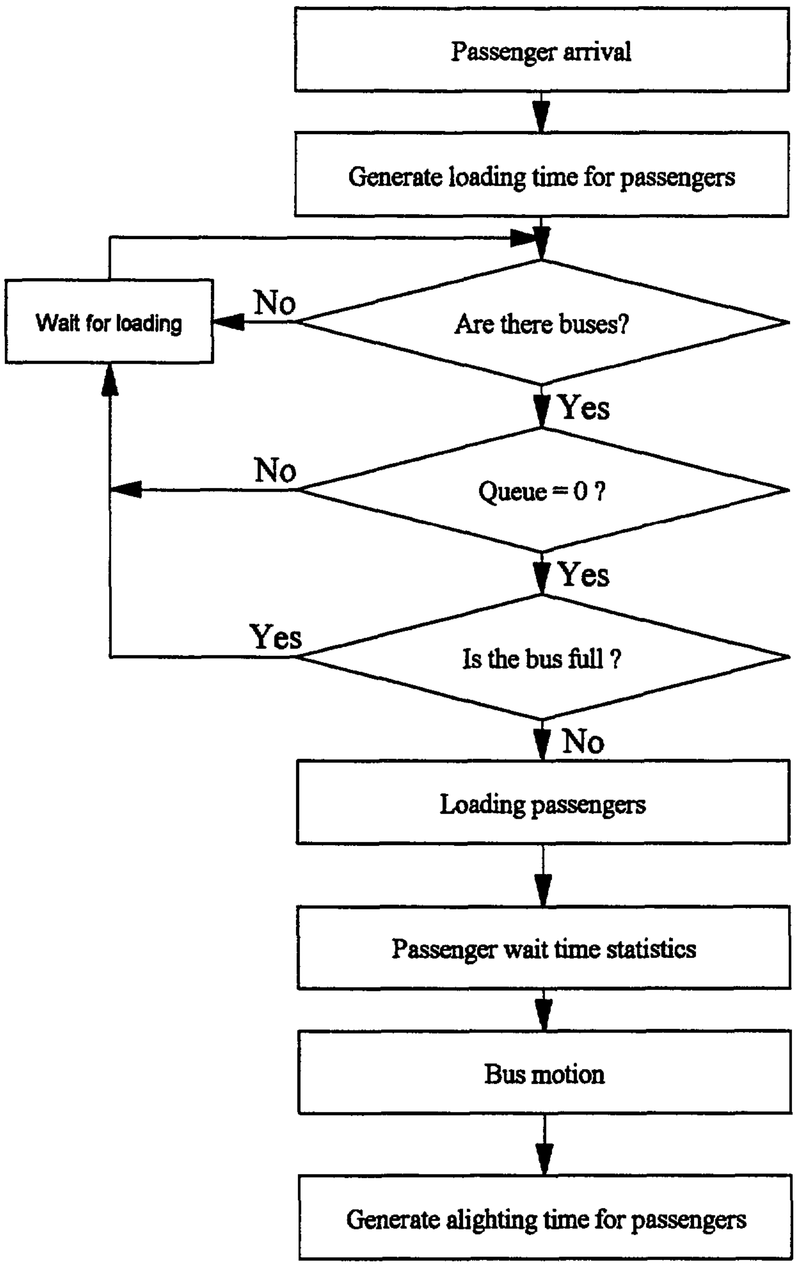

Figure 3-1 Flow chart for passenger stream

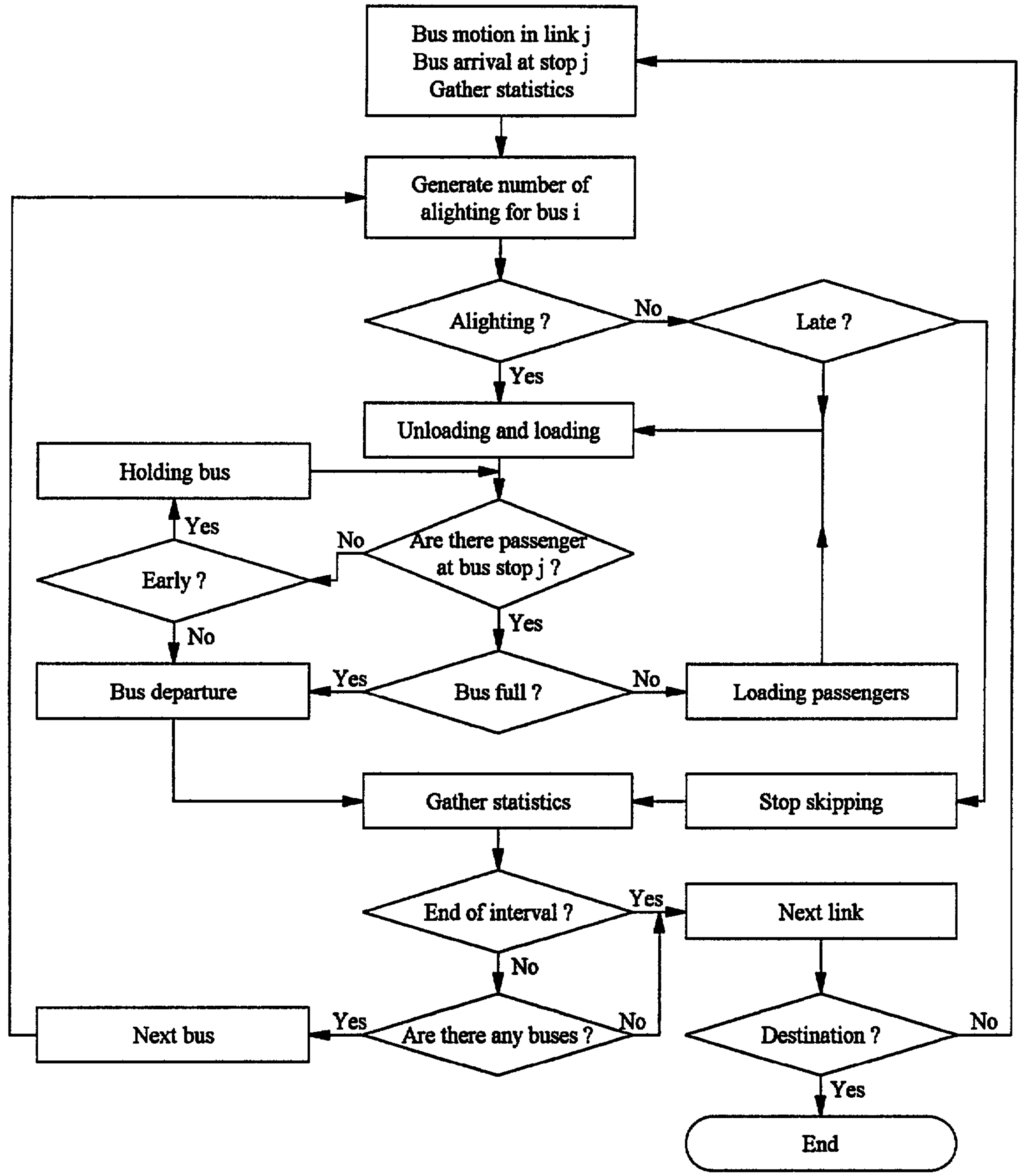

Figure 3-2 Flow chart for bus stream

3. Passenger arrival is a Poison process, i.e., the number of arrivals per unit time is Poison-distributed. This implies that the interarrival time between two successive passengers is exponentially distributed.

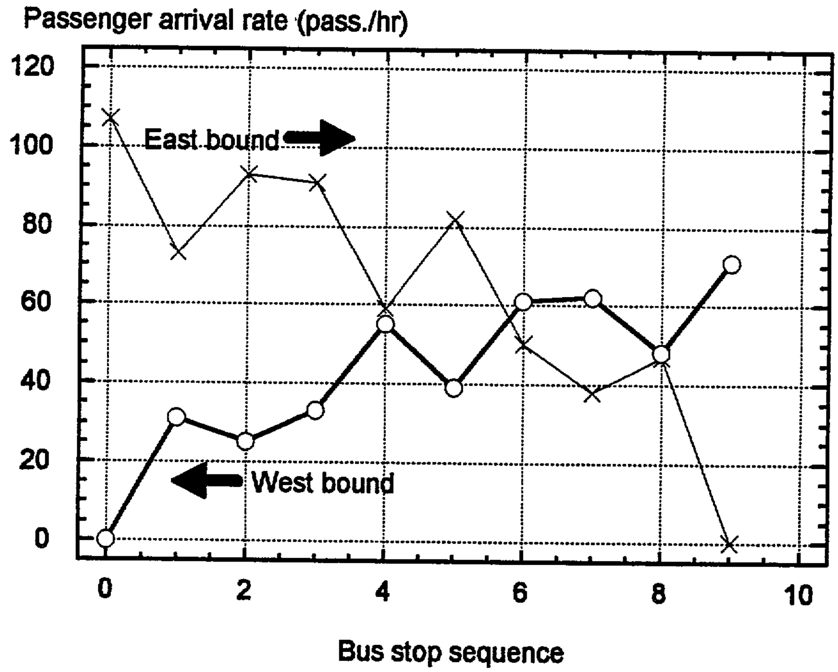

Figure 3-3 Passenger arrival rate along bus route

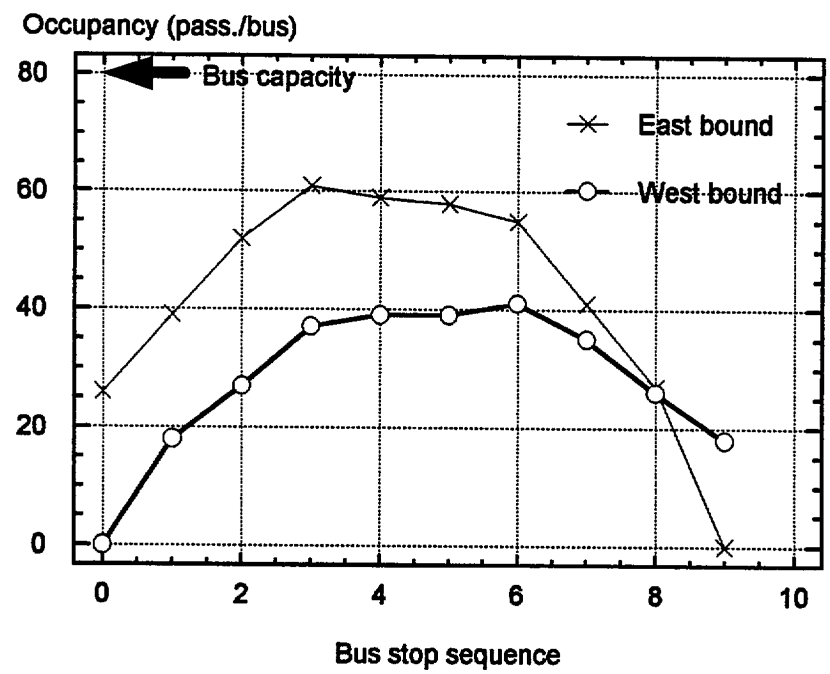

Figure 3-4 Bus occupancy along bus route

4. Bus movements are influenced by traffic fluctuation. The bus travel time used in each link (between two adjacent bus stops) is extracted from the results of TRAF-NETSIM simulation on the assumed network. At the signalized intersections, buses may join the vehicle queues and be delayed. **In** this chapter, it is assumed that the bus route traverses an urban area with unsignaliied intersections. It is further assumed that this bus route follows the main streets, and stop and yield signs control only the minor streets. Thus, the buses can travel through each intersection without any intersection approach delay. In Chapter 5, the bus movement on a bus route with signalized intersections will be modeled, and the effect of fixed signal timing and adaptive signal control at intersections on bus operation will be discussed.

5. Loading and unloading time are independent of each other. Loading time and unloading time for each passenger is 2-Erlang distributed, as proposed by Kraft [54]. In our model, the average loading time and alighting time are taken as 4.2 seconds and 2.1 seconds, respectively, which are close to the values suggested by Koffman [14].

A simulation experiment produces the following outputs:

(1) Waiting time (T_W): This is the time a passenger spends for waiting buses at the bus stop. It is counted from the time for a passenger to arrive at the bus stop through the time for the bus to depart from that stop. In our model, the average wait time of passengers is counted for all passengers served during the whole period of interest. Thus, the effect of a control option on the passenger wait time at downstream stops is also considered.

(2) Standard deviation of passenger wait time (s_W): For a uniform arrival assumption, the deviation of passenger wait time is $H/\sqrt{12}$. Since the assumption of Poison arrivals is made in our model, the simulated result of STD of passenger wait time may be slightly different from uniform cases.

(3) In-vehicle time of passengers (T_V): This is the time a passenger spends on the way, including bus moving time, acceleration time, deceleration time, delay due to signal

control, and the dwell time a on-board passenger spends due to bus stopping at bus stops.

(4) User time (T_u): This is the sum of waiting time and in-vehicle time of passengers. In the model, the average user time does not include the access time to and exit time from the bus system since access and exit time are not related to bus movement controls.

(5) Bus travel time (T_b): This is the total one-way travel time a bus spends from its departure from the original terminal until its arrival at the destination station. Due to the different passenger demands in the east bound and west bound directions, the bus travel times in the two directions will be counted separately. The bus round trip time is the sum of the two-way travel times and average dwell time at bus terminals. It is used to calculate the required number of buses.

(6) Average headway (h) and its standard deviation (s_h): Average headway is the mean of observed bus headway at all bus stops along the route.

(7) Average user cost per hour. This is the product of the total passenger hours per hour and value of passenger time.

(8) Average supplier cost per hour. This refers to the bus operating cost per hour. It is the product of the required bus number and value of bus time.

(9) Total operating cost per hour (C): the sum of total user cost and supplier cost per hour.

3-l-2 Description of Control Strategies at Bus Stops

In practice, holding and stop-skipping controls at bus station are often applied for improving bus service reliability. Holding control is used to deliberately slow down an early vehicle, and stop-skipping control is used to speed up a late vehicle. From the control logic, two important types of bus operation control can be distinguished. One type is focused on maintaining constant headway between successive vehicles, and the other one is oriented toward controlling vehicles to a particular schedule. The former strategy is referred to as headway-based control, and the latter one is referred to as schedule-based control. Each hind of control may be binary (all or nothing) or proportional (proportional to desired headway or pre-planned schedule), as tabulated below:

	Binary Control	Proportional Control
Headwav-based Control	BIH	PRH
Schedule-based Control	**BIS**	**PRS**

1. Headway-Based Control

The major objective of the headway-based control strategies is to maintain proper bus headways (that typically means equal headways) in order to reduce bus bunching and passenger wait time. In 1972, Osuna and Newell have derived the following expression for expected waiting time of randomly arriving passengers [13]:

$$E(W) = E(H) / 2 + Var(H) / 2E(H) \qquad (3.0)$$

where

$E(W)$ = average wait time for randomly arriving passengers,

$E(H)$ = average headway between buses, and

$Var(H)$ = variance of headway.

Thus, reducing the variance of headway will decrease the average waiting time when the frequency of buses is specified by the operator.

Two subclasses of headway-based control are Binary Headway-Based Control (BIH) and Proportional Headway-Based Control (PRH). "Binary" implies two options: full control or no control. A binary headway-based control strategy maintain each headway between tolerable bounds αH and βH with respect to a preceding vehicle. αH is the smallest allowed headway from the previous bus, which is called the early threshold or early bound. H is the pre-planned headway, and a is the holding control parameter ($0 \leq \alpha$

≤ 1). ßH is the largest allowed headway from the previous bus, which is called as late threshold or late bound. ß is the skipping control parameter ($ß \geq 1$). αH and ßH determine a control range. Specially, a = 0 represents no holding control, and positive infinite β represents no stop-skipping control. The control range should be optimized by the bus operators. Generally, the movement of a bus will not be controlled unless its trajectory is beyond the bounds.

Figure 3-5 is a time-space diagram of bus movement. In the figure, line 1 (early bound) and line 2 (late bound) define a control range for bus k+l . When the headway between the previous bus (bus k) and the arriving bus (bus k+l) is less than αH, the arriving bus will be held until the headway is up to αH. If the actual headway is greater than ßH, the arriving bus may be instructed to skip the stop. If bus k+l is within the control range, it will not be controlled at all.

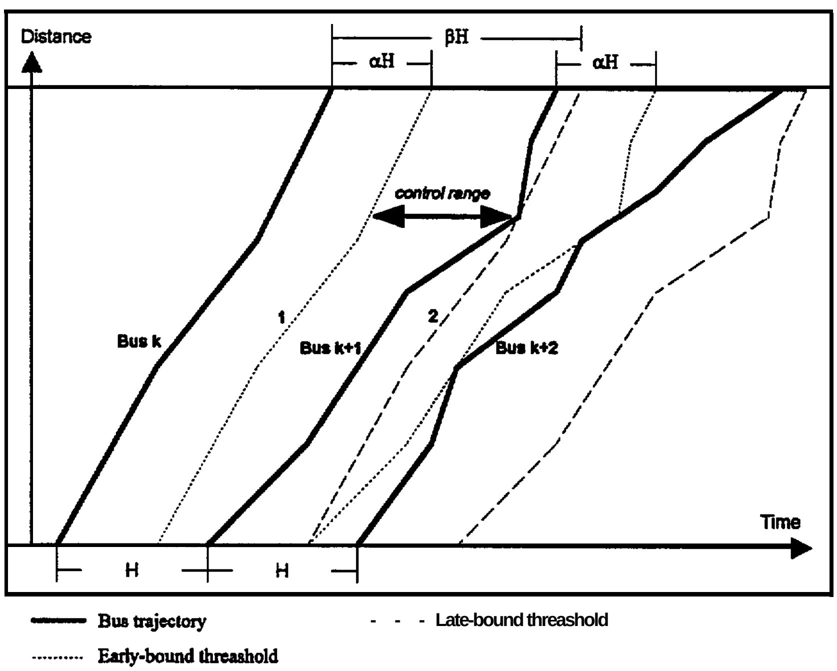

Figure 3-5 Headway-based control

For Proportional Headway-Based Control, the strength of holding control parameters is proportional to the deviation from a pre-planned headway H, and no rigid early bounds are applied. When a bus is closer to the previous bus than the pre-planned headway, it will be held for a certain time which is the product of the deviation from the pre-planned headway and a given holding ratio α_1 ($0 \leq \alpha_1 \leq 1$). Thus, the more a bus deviates from the expected trajectory, the more time it will be held. The larger the specified holding ratio, the more held time. For instance, in figure 3-5, the bus k+l has an early deviation from its pre-planned headway, that is, $(t^j_k + H - t^j_{k+l}) > 0$. Here, t^j_k is the departure time of bus k at stop j, and t^j_{k+1} is the time for bus k+l at stop j to be ready to depart, and again H is the pre-planned headway, the bus should be held for $\alpha_1(t^j_k + H - t^j_{k+1})$. If the pre-planned headway is 5 minutes, and the actual headway is 3 minutes, there is 2-minute early deviation for this bus. When 0.8 of holding rate is used, the holding time for the early bus should be 0.8 x (5 - 3) = 1.6 min.

It should be noted that even if the bus k+l at bus stop j is within the control range consisting of line 1 and line 2, the controls are still needed, because bus k+l has deviated fiom its pre-planned headway H. Hence, proportional control will hold all early buses. Its objective for this kind of control is to pull early buses gradually back to their desirable trajectory. This is the difference between binary and proportional controls. With proportional control, the rigid late bound is still needed since proportional skipping ratio doesn’t make sense. Therefore, the proportional control is the combination of proportional holding control and binary stop-skipping control.

Both binary headway-based control and proportional headway-based control correct the trajectory of a bus according to its location relative to the previous bus. Thus, the departure information of the previous bus at the bus stop should be collected and recorded, and then transmitted to the following bus. Thus, the pre-planned bus schedule plays no role under this type of control policies.

2. Schedule-Based Control

Schedule-based strategies control buses toward keeping the original schedule instead of maintaining a desired headway. Therefore, the location of the previous bus is irrelevant. In schedule-based control strategies. Buses are controlled to adhere to their own schedule regardless of how much bus bunching occurs. Nevertheless, when bus movements are close to the schedule, bus bunching will be reduced. Therefore, schedule-based controls can indirectly maintain regularity of bus headway.

Similarly, schedule-based controls can be classified into Binary Schedule-Based Control (BIS) and Proportional Schedule-Based Control (PRS). For binary schedule-based control, again, "binary" implies two option: Ml control or no control. Given a pair of tolerable deviation parameters from schedule, $\alpha'H$ (early tolerable deviation value or early bound) and $\beta'H$ (late tolerable deviation value or late bound), in which a' (≥ 0) is the holding control parameter, $\beta' (\geq 0)$ is the skipping control parameter, when a bus is more than $\alpha'H$ minutes ahead of the planned schedule (earlier than the early bound), it will be held until its actual deviation is less than or equal to the tolerable deviation from schedule (see figure 3-6). If a bus is more than $\beta'H$ minutes behind the schedule (later than the late bound), it may be instructed to skip bus stops until it returns within the given control range. If the bus is within the control range consisting of the early bound (line 1) and the late bound (line 2), it is not controlled at all.

Binary schedule-based control can be implemented easily, because bus drivers operate their vehicles only according to the planned schedule and the given tolerable deviation values. The direct objective of the control strategy is to increase on-time performance of bus operation in order to prevent bus bunching.

Similarly to proportional headway-based control, proportional schedule-based control holds all early buses for a certain time according to the given holding ratio $\alpha'_1 (0 \leq \alpha'_1 \leq 1)$. Unlike proportional headway-based control, its holding time is computed according to the deviation of the bus from its schedule, instead of its headway with respect to the preceding bus. For instance, bus k+l located at stop j is ahead of its schedule ($t^j_{k+l} < T^j_{k+1}$), in which t^j_{k+1} is the time for bus k+l to be ready to depart from stop j, and T^j_{k+l} is the schedule of bus k+l at stop j. Thus, the bus will be held for $\alpha'_1(T^j_{k+1} - t^j_{k+l})$.

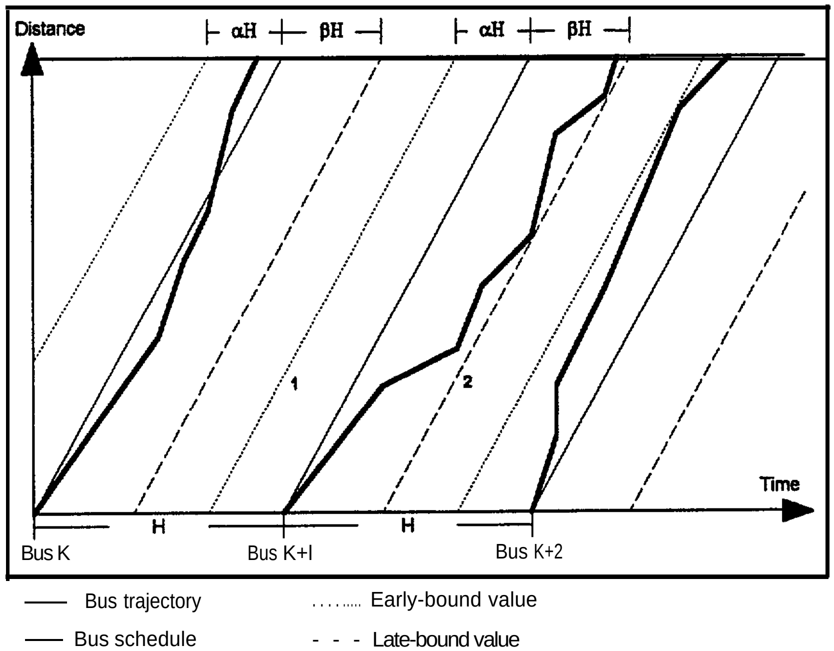

Figure 3-6 Schedule-based control

This chapter focuses on the analysis and comparison of headway-based control (BIH), proportional headway-based control, schedule-based control, and uncontrolled operation. For these strategies, the controllable variables are holding and stop-skipping control parameters.

3-2 Analysis of Experimental Results

In this section, bus operation performance under the Binary Headway-based (BIH) and Binary Schedule-based (BIS) control strategies are analyzed. Analysis of Proportional Headway-based control (PRH) and Proportional Schedule-based control (PRS) are not given in this research report since their optimal objective function . value is equivalent to that of BIH and BIS respectively [55].

3-2- 1 Holding Control

Holding controls are used to delay bus movement deliberately when a bus is ahead of the planned schedule or too close to the previous bus, Holding control can produce the following results:

1. Tight controls can significantly reduce the headway variance (figure 3-7). This implies that under tighter holding controls, the headways are distributed more evenly. Here, tight control means that buses are brought more closely to the schedule or desired headway. It should be noted that in the figures, the values of holding control parameters are not given because the holding control parameters for the different control strategies involve different concepts and measurements. The horizontal axis just shows the directions of looser control and tighter control. The lefl end of the scale in the figure represents the uncontrolled option, and the right end represents that early buses are held until the schedule or the pre-planned headway is satisfied. The two curves on figure 3-7 show that headway variance can be improved more through the headway-based control than through the schedule-based control. This is because under headway-based strategies bus movements are controlled toward headway regularity, while under schedule-based strategies bus movements are oriented toward on-time performance.
2. Tighter holding control can reduce the average wait time of passengers (figure 3-8 and 3-9). It can be observed that the headway-based strategy can yield a lower wait time than the schedule-based strategy. This is because the headway-based control can reach a smaller headway variance than the schedule-based control. Generally, wait time decreases as headways are more equal (i.e. as headway variance decreases).

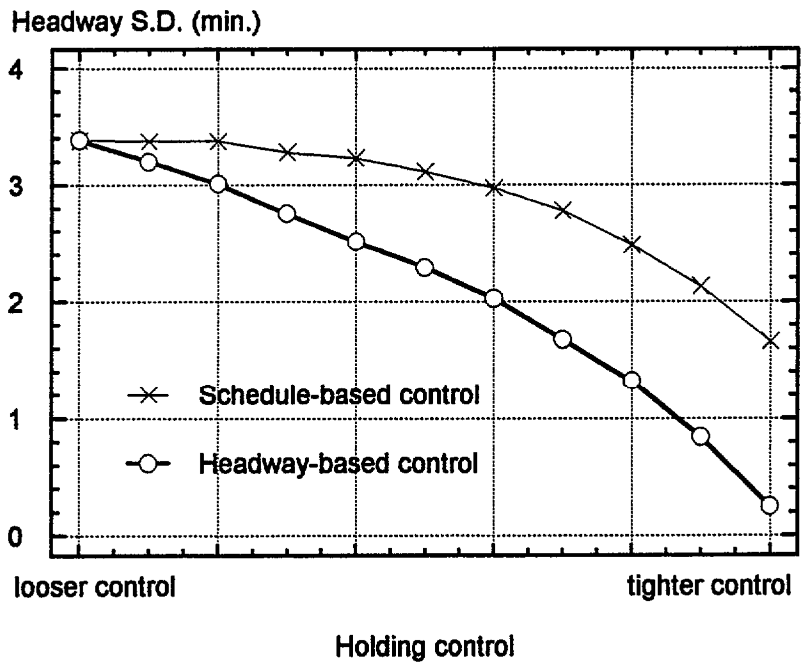

Figure 3-7 Headway standard deviation vs. holding control

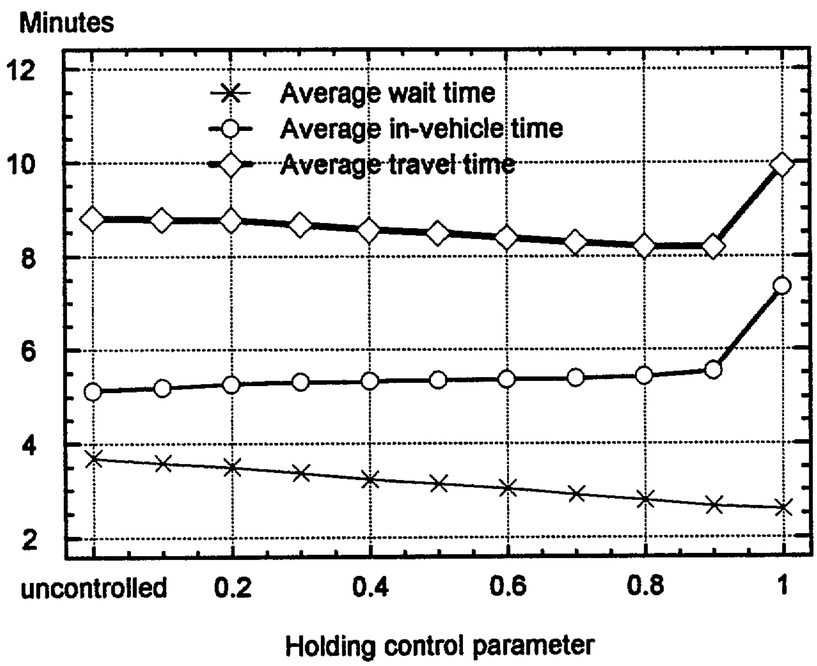

Figure 3-8 Headway-based holding control

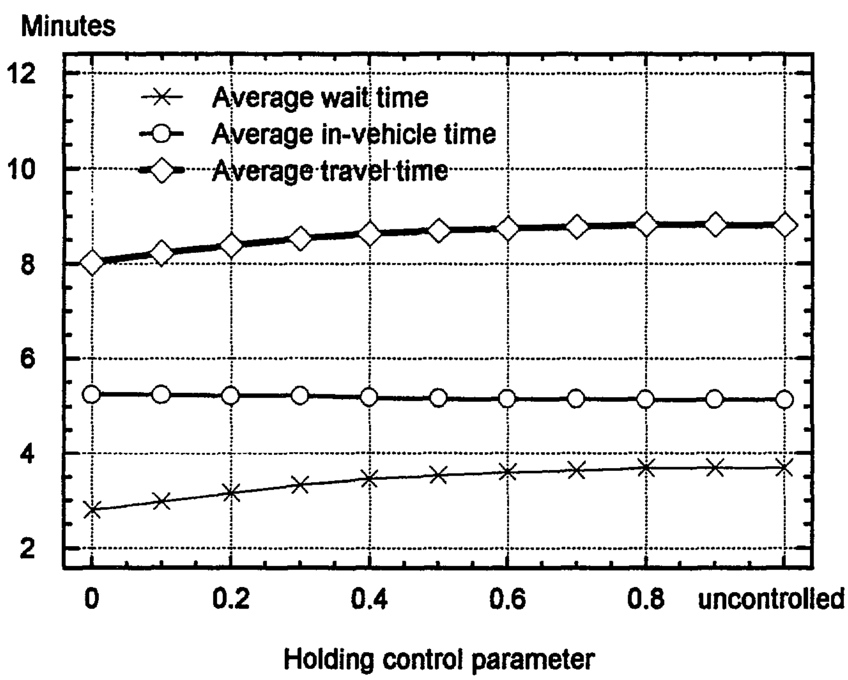

Figure 3-9 Schedule-based holding control

3. Tighter holding controls will increase in-vehicle time of passengers (figure 3-8 and 3-9). The in-vehicle time includes moving time and dwell time. In general, the moving time depends on the traffic conditions. It is not related to our control options implemented at bus stops. Thus, we can conclude that tighter holding control results in more dwell time for on-board passengers. The average travel time of passengers consists of average wait time and average in-vehicle time. It should be noted that average travel time of passengers under tighter schedule-based controls tends to decrease, while the average travel time under headway-based controls seems convex. Under headway-based strategies, a tighter holding control reduces the user time. However, when the holding control value exceeds some point, the user time gets worse than under uncontrolled operation. This is because the increase of in-vehicle time exceeds the decrease of wait time.

4. Tighter holding controls increase bus travel time. In particular, tighter holding under headway-based controls has significant effects on bus travel time. This is because tighter holding control reduces the chance of early bus departures. However, when the previous bus is late, the following bus will be shifted right (i.e. delayed deliberately) to maintain the desired headway. Thus, the trajectories of buses will tend to slope rightward. Obviously, the bus travel time under headway-based controls will also increase, as shown in figure 3-10. Nevertheless, under schedule-based controls, buses are dispatched according to their schedule. The location of the previous bus does not impact the dispatching decision for the following buses.

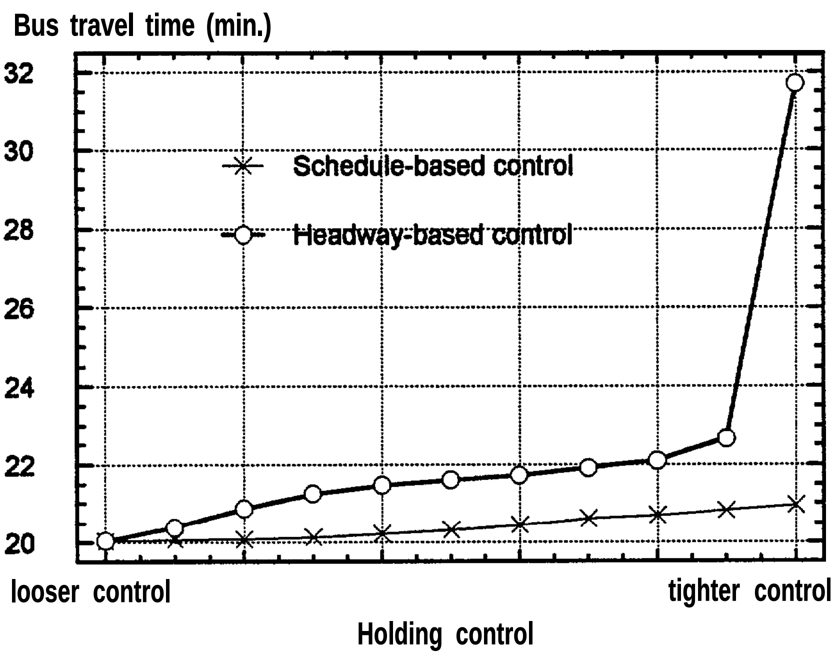

Figure 3-10 Bus travel time with holding control

3-2-2 Stop-Skipping Control

The objective of stop-shipping controls is to prevent bus lateness, which is another major cause of bus bunching. The stop-skipping option can be used with both headway-based controls and schedule-based controls to avoid bus lateness. Nevertheless, the feasibility of adopting stop-skipping control is often doubtful. Skipping stops would increase bus travel speed (line-haul speed), and decrease the in-vehicle time of on-board passengers and the wait time of passengers at downstream stops. However, its expense is increase of the wait time for those passengers passed by. Moreover, passengers watching a bus going past without stopping might be discouraged from using the bus system.

This experimental result shows that the passenger wait time increases significantly with tight stop-skipping controls, as shown in figures 3-11 and 3-12. This implies that a tight skipping control leaves more passengers waiting for the following buses, even though it can significantly improve the regularity of bus movements, as it is shown in figure 3-13 and 3-14 that tighter skipping controls can reduce the deviation of bus headways.

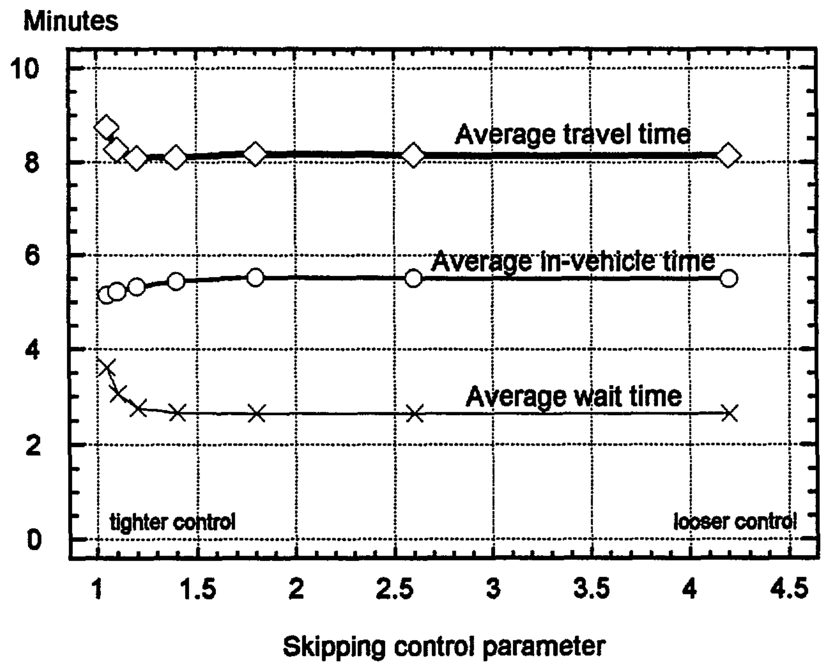

Figure 3-11 Service time for headway-based skipping control

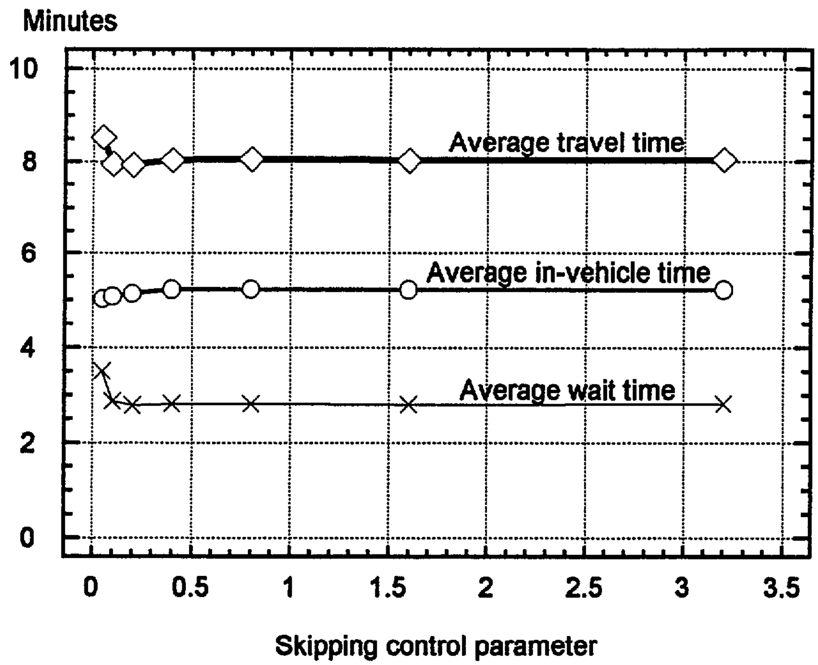

Figure 3-12 Service time for schedule-based skipping control

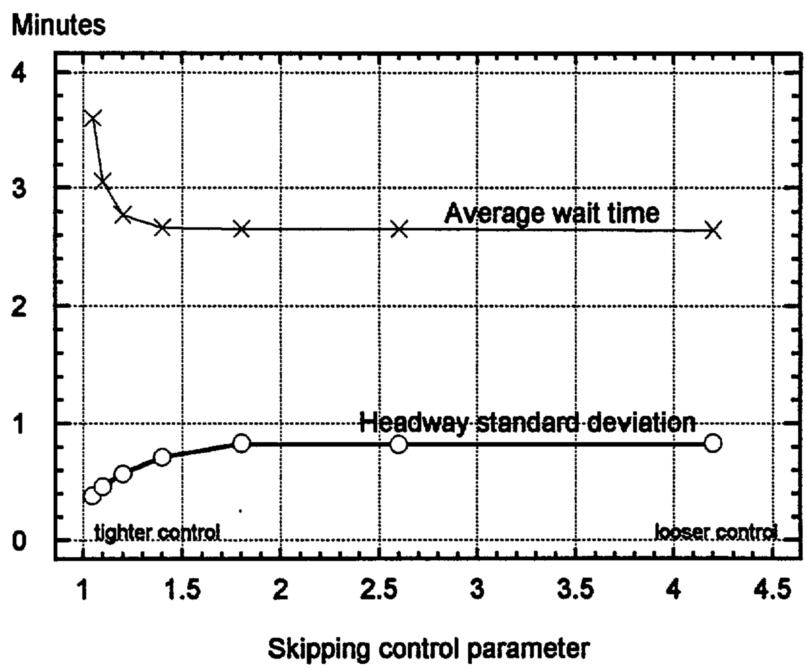

Figure 3-13 Bus headway deviation for headway-based skipping control

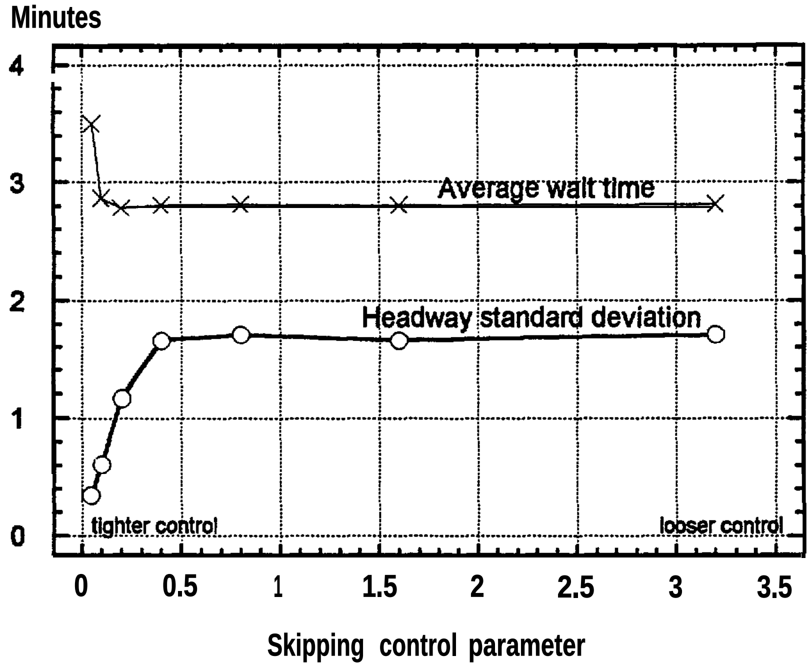

Figure 3-14 Bus headway deviation for schedule-based shipping control

This result also shows that tight stop-shipping control can reduce slightly the in-vehicle time. Nevertheless, passenger travel time, the sum of wait time and in-vehicle time, increases as stop-shipping control parameters close to the pre-planned headway or schedule, Hence, tight slopping controls should be avoid. It can be seen that the curves of wait time and travel time decrease and become flat as the shipping control loosens. This implies the stop-shipping control may not be implemented.

3-3 Optimization and Evaluation of Control Strategies

The previous analysis explored individual effects of holding and shipping controls on bus movement and performance. Naturally, a pair of control values (i.e. the early bound and late bound) should also be considered simultaneously. Holding control values and shipping control values, which are selected as decision-making variables, form continuous parameter spaces. The purpose of this section is to (1) construct an objective function, and (2) search for good combinations of the two control values that minimize the specified objectives.

3-3-l Objective Function and Optimal Combination of Control Parameters

Three objective functions are often particular to bus operators or bus users. They are average wait time of passengers, average user cost, and total cost. Decision-makers can select one or a combination of these objectives according to their actual requirement and preferences. The total cost function is formulated as follows:

$$C=Q(c_w T_w+c_v T_v)/60+(T_b+T_d)c_b f \quad (3.2)$$

where

c: total cost ($/hr.)

Q: average ridership per hour (passengers/hour)

T_w: average wait time of passengers (minutes)

T_v: average in-vehicle time of passengers (minutes)

c_w: value of passenger wait time ($/hour)

c_v: value of passenger in-vehicle time ($/hour)

T_b: bus average round-trip time (minutes)

T_d: origin and destination terminal layover time (minutes)

f planned frequency of bus (buses/hour)

c_b: bus operating cost rate ($/hour)

In the cost functions, T_w, T_v, and T_b vary with control options. Their values will be obtained through this simulation model.

A numerical example is presented to explore the effect of control options on user cost, supplier cost, and total cost. In the example, the following parameters are used:

value of passenger wait time = $16/hour;

value of passenger in-vehicle time = $8/hour;

bus operating cost rate = \$50/hour/bus;

origin and destination terminal layover time = 10 minutes.

Figure 3-15 displays the total cost under headway-based controls. Four holding control options and seven slopping control options are combined into 28 candidates. These costs plot a convex cost response surface. It can be observed that when the holding control parameter is 0.9 (the tolerable smallest headway = 0.9x5 = 4.5 min.) and skipping control parameter is 1.4 (the tolerable largest headway = 1.4x5 = 7 min.), the minimum total cost can be obtained. The result is consistent with the previous analysis: for headway-based controls, overly tight holding and slopping controls are undesirable,

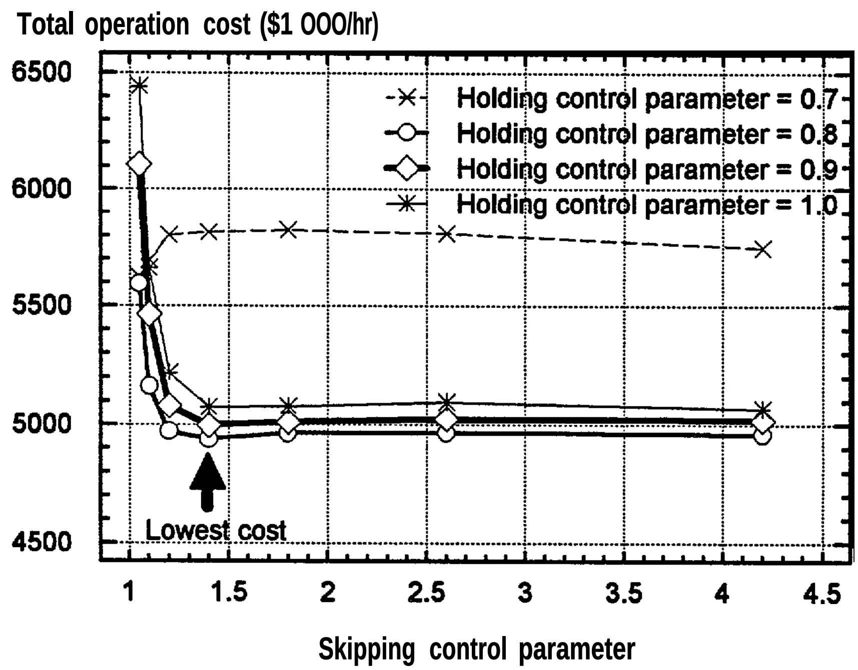

Figure 3-15 Total operation cost for headway-based control

For schedule-based controls, the tight holding control increases in-vehicle time and bus travel time. But, their increases are not significant. However, tight holding control is quite beneficial in decreasing wait time and corresponding total cost, unless the bus cost is the larger fraction of the total cost than the passenger cost. In this example, the fraction of user cost is much more than that of supplier cost. Therefore, the a tight holding control (until the pre-planned schedule) is applied. From figure 3-16, it can be seen that the skipping control value should be about 0.2 (i.e. 20%) of headway. Overly tight skipping controls increase the total cost.

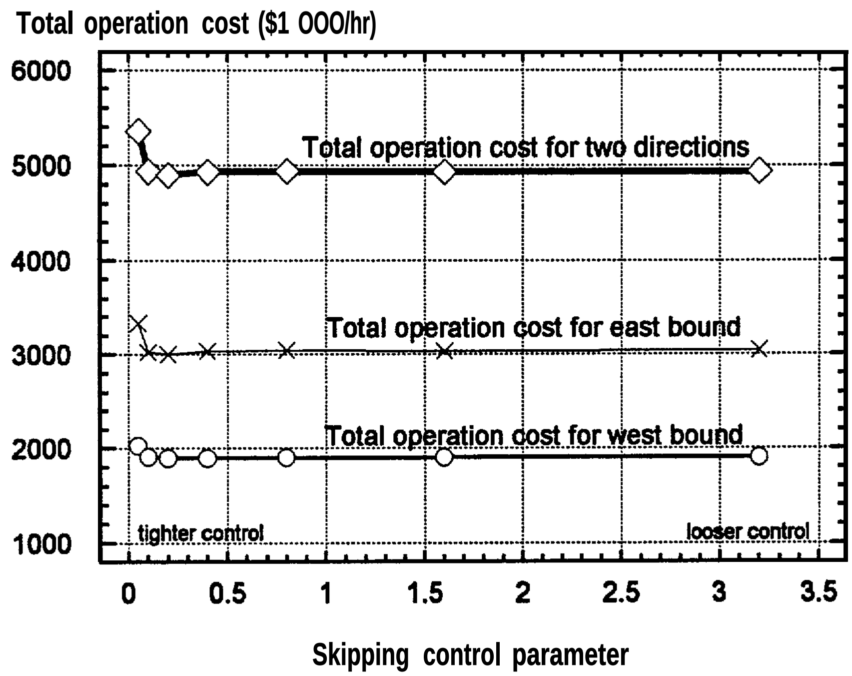

Figure 3-16 Total operation cost for schedule-based control

3-3-2 Comparison of Headway-based and Schedule-based Strategies

Based on the simulation results, the comparison between the headway-based controls and schedule-based controls is summarized in the following table. In the table, there are six decision-making objectives. The value in the table shows the best level the headway-based and schedule-based control strategies can reach with different holding and stop-slopping control parameter combinations. For minimum average wait time, minimum average user time, minimum travel time of buses, and minimum headway deviation, the east bound and west bound values are separated by a slash "/". In this table, the operation performance under no control also is given in order to reflect the improvement under controls.

Table 3-1 Comparison of the best headway- and schedule-based controls

Objective	Headway-based control	Schedule-based control
Minimum total cost ($/hour)	4938	4891
Minimum user cost ($/hour)	4436	4411
Minimum average wait time (minutes)	2.59 / 2.52	2.79 / 2.65
Minimum average user time (minutes)	8.09 / 7.23	7.92 / 7.12
Minimum average travel time of buses (minutes)	19.8 / 16.6	19.8 / 16.6
Min. standard deviation of headway (minutes)	0.19 / 0.10	0.34 / 0.33

-Key: East Bound / West Bound

The result shows that at this specified 5 minute headway case, if total cost, user cost, or user time is selected as decision-making objective, schedule-based controls are preferred. If wait time or headway standard deviation is chosen as objective, headway-based strategies are preferred. Headway-based strategies can reach better performance in wait time and regularity of bus movement than schedule-based strategies, however, they have a higher total cost. It implies that greater regularity of bus services may not always be consistent with lower passenger cost. Therefore, the deviation of headway should not be taken as a unique decision-making criterion.

Chapter 4 Adaptive Signal Control for Transit Buses

Preemption strategies are intended to provide relatively high priority to transit buses lower delays through signalized intersections, thus improving overall travel speed and reducing travel time. Most control models with such strategies have focused on reducing the travel times for mixed traffic rather than particular other modes. Models with more complex features may possibly treat traffic situations (e.g. priority request of bus movements) more accurately. However, they lose capability of promptly responding such situations within a short time period. Conversely, models with simpler logic could deal with such situations quickly but fail to control over the entire traffic judiciously.

With the above limitations, two key issues arise in developing a signal preemption model for buses. First, how should a model account for the traffic operating cost at an intersection when giving priority to buses? Second, is the model simple enough to quickly evaluate possible situations and make decisions through its optimization logic? To deal with these issues, the following aspects of control model are considered:

1. Design of signal base plan.
2. Formulation of traffic operating cost for signals.
3. Logic of signal preemption for buses.
4. Timing optimization for bus priority.

Each of the above modules, with its own preliminary assumptions, is characterized in the following sections. Some control parameters of the model, such as bus service headways and signal cycles, are also analyzed using simulation in the last section.

4.1 Fundamentals of Signal Timing Plan

A signal base plan is a timing scheme designed by referring a long term history of traffic demands at the controlled intersection. Practically, for a n-phase fixed time signal, the cycle length should be pre-determined before it is split into phases. Various ways for designing signal cycles have been developed, such as the minimum cycle method, cycle performance method, minimum total delay method, and Webster's model [56, 57, 58, 59]. The best known method for designing signal cycles is Webster's model which estimated an optimal cycle considering minimum delay criteria. Its cycle formula is as follows

$$\text{Optimal cycle length } C_o = \frac{1.5L+5}{1-\sum y_{ci}} \quad \text{......} \quad (4.1)$$

where

L = Total lost time in one cycle, in seconds

y_{ci} = Flow ratio $y_{ci} = q_{ci}/s_{ci}$ for critical approach ci

The numerator of Webster's C_o formula has been modified to fit different traffic characteristics in some countries, such as $(1.4L + 4)$ in Austria and $(1.4 + k)\ L + 6$ in Australia (where k is the lost weight constant for stopped vehicles) [60, 61]. By considering such formulas, A function $C = f(L, X_C, y_{Ci})$ is used in this study. Given that the average flow rate and the saturation flow for critical movement ci are q_{ci} and s_{ci}, a n-phase signal cycle for the controlled intersection can be shown as

$$C = \frac{X_c L}{X_c - \sum y_{ci}} \quad \text{(critical movement } i = 1, 2, ..., n) \quad \text{......} \quad (4.2)$$

where

C = Cycle length in seconds

L = Total lost time in seconds/cycle ($L = 3.0 \times n$ seconds per cycle in this study).

X_c = Critical ratio of the intersection. The 1985 HCM recommends $X_c = 0.95$ to accommodate higher flow fluctuations [62, 63, 64].

The phase splits of cycle C can be calculated based on $y_{ci} = q_{ci}/s_{ci}$ ratios ($i = 1 \sim n$). The critical ratio X_c indicates the proportion of available capacity that is being utilized by vehicles in critical movements. If this ratio exceeds 1.0, one or more of the critical movements will be over-saturated. A ratio of less than 1.0 is an indication that the intersection design, cycle, and phase plan are adequate for the existing or projected traffic demand, assuming that green splits are proportionally assigned [62]. For a n-phase signal, three requirements are followed in this study:

(1) Any two effective green phase durations g_i and g_j should have the ratio

$$\frac{g_i}{g_j} = \frac{q_{ci}s_{cj}}{q_{cj}s_{ci}} \quad \text{.......} \quad (4.3)$$

(2) Any effective red phase duration r_k, Min $r_k \le r_k \le$ Max $r_k = C(1 - \frac{q_{ck}}{s_{ck}})$ .... (4.4)

(3) For any phase k, $g_k = \sum_{i=1}^{n} r_i - (n-2)C - L$, and $r_k = \sum_{i=1}^{n} g_i + L$ (where $i \neq k$) ..... (4.5)

In a 2-phase signal design (n = 2), requirement (3) can be simplified into $g_1 = r_2 - L$, and $r_1 = g_2 + L$. The basic concept is shown in figure 4-1.

4.2 Traffic Operating Cost at Signals

Traffic demands at an intersection may vary in any time interval. The time interval is a time step during which the flow rate on approach i, q_i, saturation flow s_i, and some associated traffic performance measures are computed. The traffic pattern on each approach during any time interval is regarded as uniform over time. Thus, an average flow rate $\bar{q}_i$, rather than q_i, can be used to develop a base signal timing plan when the time interval becomes extremely long. Based on the timing plan, the traffic performance measured during the time period can also be expressed on an average basis (e.g. cost per unit of time, or delay per vehicle).

In a 2-phase signal design, figure 4-2 shows how the traffic performance (inside the shaded areas) within one cycle is measured. Instead of step flow functions, the plot uses continuous cumulative flow functions q_i (or q_j) and s_i (or s_j) for approach i on major street (or approach j on minor street). The shaded areas formed by such flow functions can be applied to many traffic performance calculations. Three of such performance measures used in this study are described as follows:

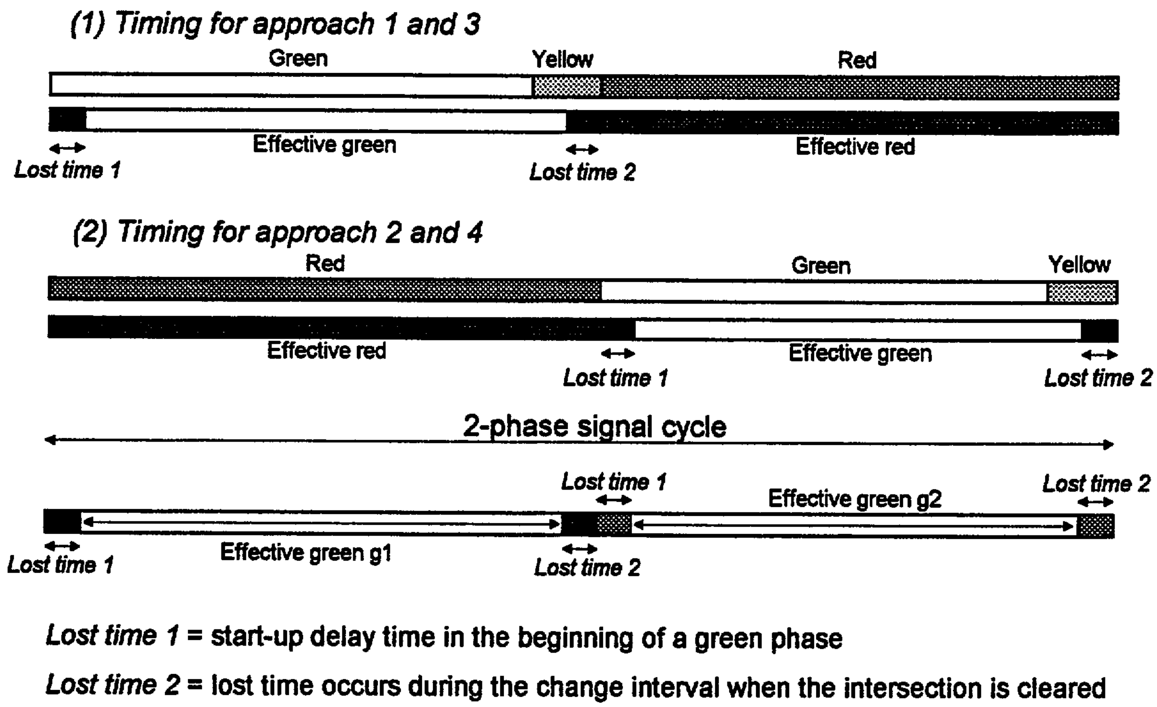

Figure 4-1 Illustration of a 2-phase signal timing

4.2.1 Passenger Car Delay

One of the most important MOE's in traffic studies is the delay to vehicles and motorists in the system. Excessive delay at a signal reflects inefficient timing setting and directly incurs time costs to motorists during idling. Thus, appropriate traffic signal setting can efficiently help smoothing traffic flow through a road network with minimum delay and stoppage.

Given that the signal phase plan is well designed to handle traffic without demand exceed capacity, the total delay (T_d) per cycle can be derived by referring the shaded areas in figure 4-2. For each intersection approach i (or j), the corresponding shaded area represents an expectation of total vehicle-seconds incurred in one complete cycle k (C_k). Therefore, the expected T_d for a 2-phase signal can be calculated by

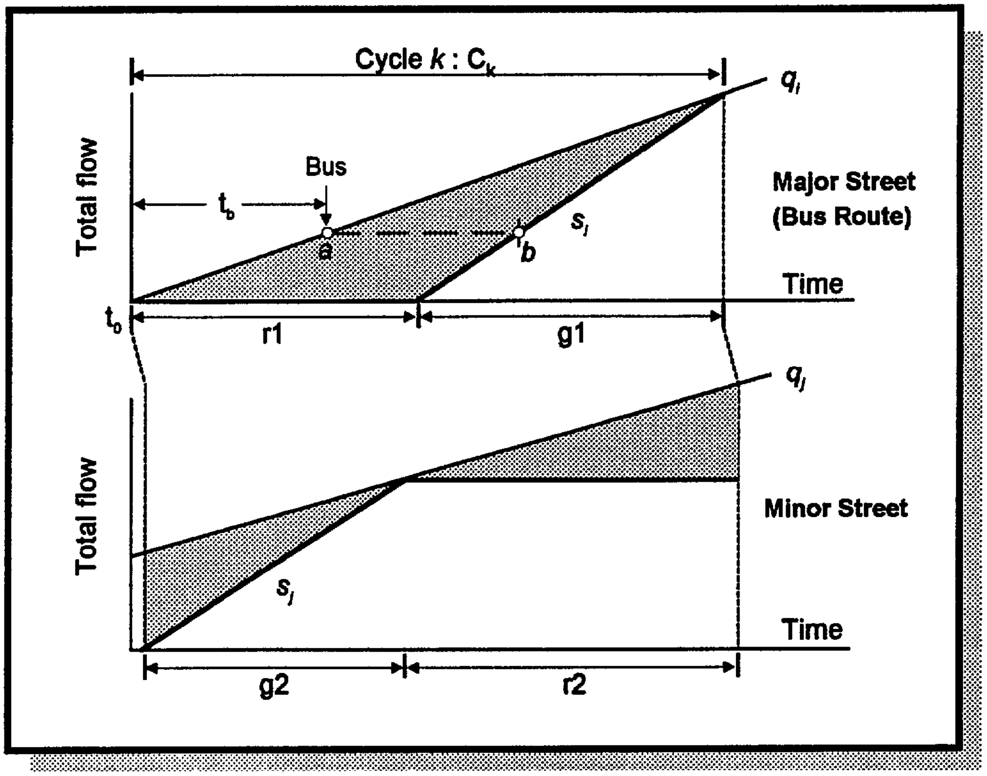

Figure 4-2 Performance measures in a signal cycle

$$T_d = \frac{1}{2}[r_1^2(\frac{\overline{q}_1 s_1}{s_1 - \overline{q}_1} + \frac{\overline{q}_3 s_3}{s_3 - \overline{q}_3}) + r_2^2(\frac{\overline{q}_2 s_2}{s_2 - \overline{q}_2} + \frac{\overline{q}_4 s_4}{s_4 - \overline{q}_4})]$$

$$= \frac{1}{2}[r_1^2(z_1 + z_3) + (C_k + L - r_1)^2(z_2 + z_4)] \quad \text{...} \quad (4.6)$$

where

T_d = Total delay per cycle, in veh-seconds/cycle

r_1 (r_2) = Effective red time for critical movement *c1* (*c2*). $r_1 = C - g_1$, in seconds.

$\overline{q}_i$= Average flow rate on approach i (i = 1 ~ 4), in veh/sec. Flows on approach i = 1 and 3 are opposing traffic on the same street. Also, flows on approaches 2 and 4 are opposing traffic on the cross street.

s_i = Saturation flow rate for approach i (i = 1 ~ 4), in veh/sec.

$z_i = (\overline{q}_i s_i)/(s_i - \overline{q}_i)$, where approach i = 1 ~ 4.

The total delay per unit of time (AT_d, in veh-sec/sec) can also be derived as

$$AT_d = \frac{1}{C_k}(T_d) \qquad (4.7)$$

To illustrate the calculations, the minimal feasible cycle $C_k = 40$ and $g_1 = 18$ seconds are obtained using equations (4.2) and (4.3) in section 4.1 with $\bar{q}_1$ ($\bar{q}_3$) = 900 vph, $\bar{q}_2$ ($\bar{q}_4$) = 800 vph, and s_i (for $\forall i$) = 2000 vph. The total delays AT_d with cycle lengths ranging from 40 to 120 seconds are plotted in figure 4-3.

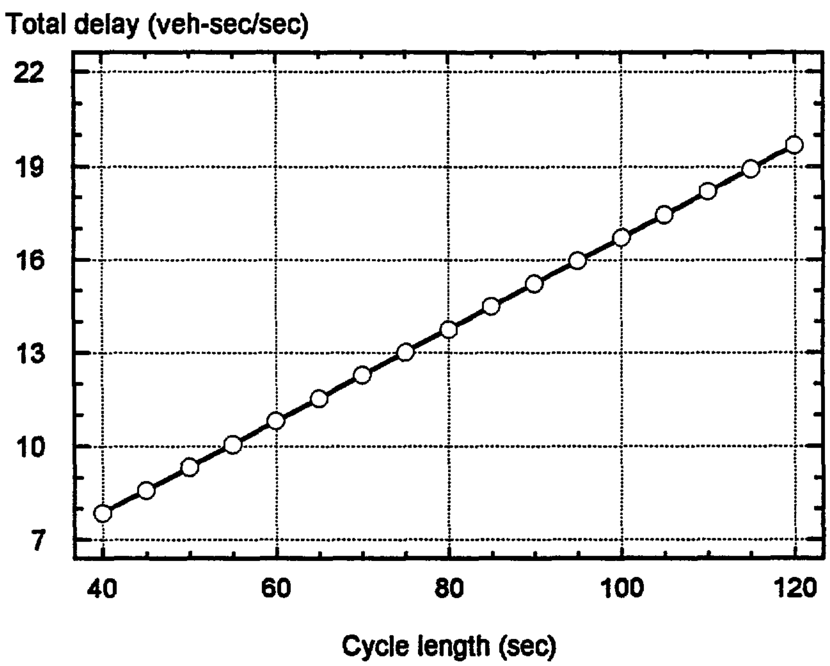

Figure 4-3 Plot of total delay vs. cycle length

4.2.2 Number of Vehicle Stops

The number of vehicle stops is a major cause of fuel consumption. From figure 4-2, the total number of stops per cycle (T_s) and the total number of stops per unit of time (AT_s) for a 2-phase signal can be derived as

$$T_s = r_1(z_1 + z_3) + r_2(z_2 + z_4)$$

$$= r_1(z_1 + z_3) + (C_k + L - r_1)(z_2 + z_4) \quad \text{.......................} \quad (4.8)$$

$$AT_s = \frac{1}{C_k}(T_s) \quad \text{.......................} \quad (4.9)$$

With the same $\bar{q}_i$ and s_i data applied in the delay measure, the total stops AT_s with cycle lengths C_k ranging from 40 to 120 seconds are plotted in figure 4-4.

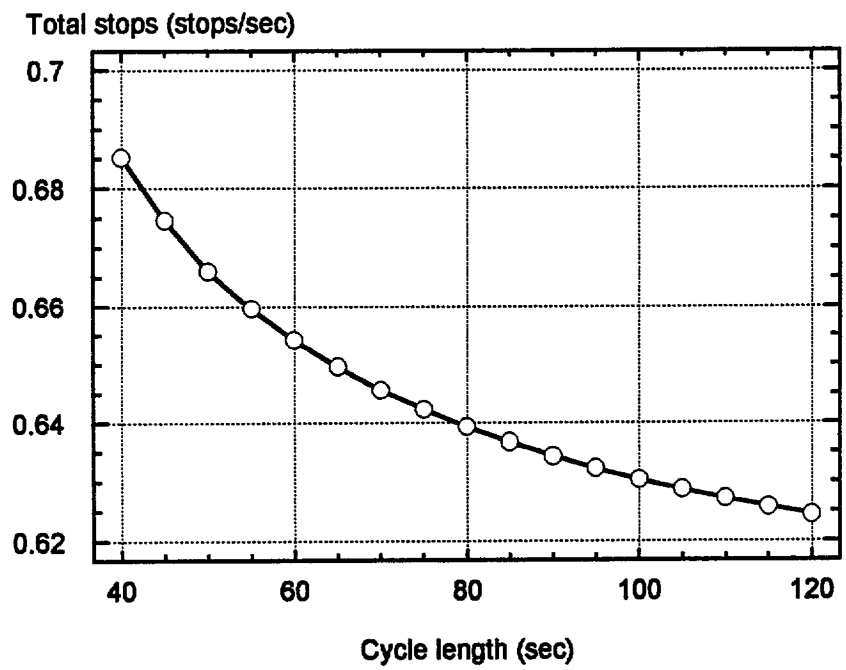

Figure 4-4 Plot of vehicle stops vs. cycle length

4.2.3 Delay to Transit Buses

Delay to a bus typically affects far more on-board passengers than delay to a car. The situation would worsen when a loaded bus is delayed by a signal. Therefore, the expected delay to buses at signals is critical for estimating the total on-board passenger delay.

For a non-saturated traffic condition, figure 4-2 can also be used to estimate bus delay (e.g. $\overline{ab}$) if the bus arrival time falls in the subject cycle k. To specify this, we let a bus arrive at the signal t_b seconds after the start of cycle k (i.e., $0 \le t_b \le C_k$). Then, the expected bus delay wt_b can be estimated as

$$wt_b = r_1 + \frac{t_b(\overline{q}_i - s_i)}{s_i}, \text{ when } t_b \le r_1 + \frac{q_i r_1}{s_i - \overline{q}_i} \text{ (approach } i = 1 \text{ or } 3)$$

$$\text{Otherwise, } wt_b = 0 \quad \text{(4.10)}$$

4.2.4 Formulation of Traffic Operating Cost

By jointly taking account of the above three measures, the traffic operating cost for a signaliied intersection in a specific cycle k during which m buses are served can be given by

$$\text{TOC} = \frac{c_d(T_{d,k}) + c_s(T_{s,k}) + \sum_j^m c_{b,j}(wt_{b,j})}{C_k} \quad \text{(4.11)}$$

where

TOC Traffic operating cost per second, in \$/second

$T_{d,k}$ Total passenger car delay in cycle k, in vehicle-seconds

$T_{s,k}$ Total number of stops in cycle k, in vehicle-stops

$wt_{b,j}$ Expected wait time of bus j, in bus-seconds

m Number of buses served over the entire cycle k

C_k Duration of cycle k, in seconds

c_d Unit cost for each passenger car delay, in \$/veh/second

c_s Unit cost for each vehicle stop, in \$/veh/stop

$c_{b,j}$ Unit delay cost for bus j, in \$/bus/second. $c_{b,j}$ is defined as a function of the passengers on bus j, the average time value of passengers on bus j, and the deviation from schedule of bus j.

It should be noted that the TOC for an intersection is calculated in units of dollars per second rather than in vehicle-time or person-time. To calculate the TOC with the same values of $\bar{q}_i$ and s_i used previously, **all** unit costs must be specified in advance. In this study, both c_d and c_s values are set to be $1 **.0/veh/second** and **$10.0/veh/stop** over time. However, a few factors such as the current deviation time **from** schedule of busj, and the number of passengers on bus j could affect the unit delay cost for bus $j, c_{b,j}$. Therefore, $c_{b,j}$ could vary and then influence the TOC function over time. To demonstrate how a $c_{b,j}$ value can affect the TOC, we assume that a bus j arrives at a signal t_b seconds **after** the start of cycle k ($0 \le t_b \le C_k$). The TOC curves considering varied values of $c_{b,j}$ are shown in figure 4-5.

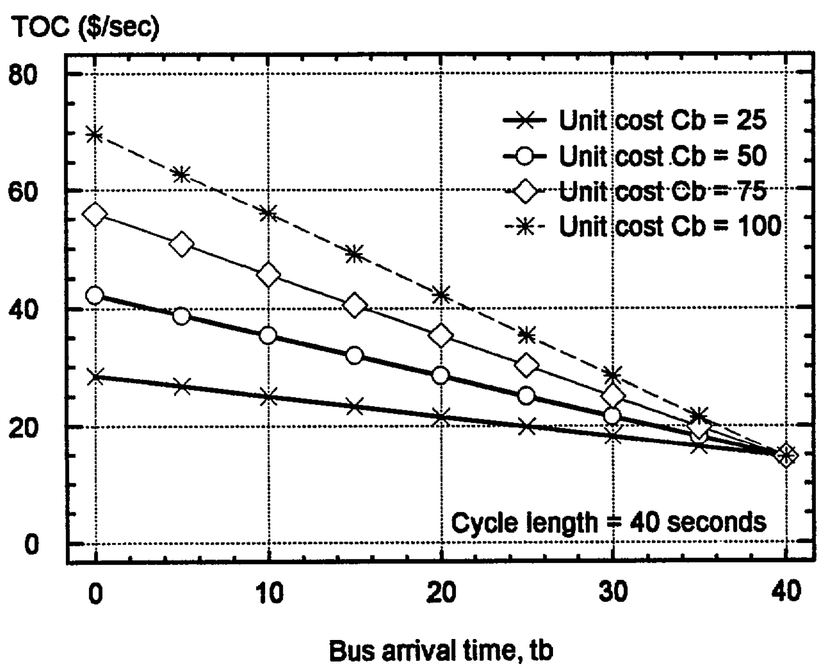

Figure 4-5 TOC vs. bus arrival time

4.3 Bus Priority Control at Signals

Transit buses can be supervised by using passage detectors, high beam transmission, license plate scanners, or some other advanced automatic vehicle location system (AVIS). An AVL system may possibly include systems with Automatic Vehicle Classification (AVC) and Automatic Vehicle Identification (AVI) functions [31]. These facilities can provide real time information when buses pass through the detection areas. For a control system without support from an AVIS, locations of the upcoming buses could be traced via the upstream detectors and/or predicted by using information from some bus stations. Basically, a detection system can provide 3- to 15-second advanced information before any bus reaches the intersections. The reliability of such prior information fully depends on where the detection facilities are located. Such advanced information could be processed with some prediction models. Then, throughout the evaluation of countermeasures, a final signal-related bus priority plan can be made before the subject buses arrive.

Some computer packages can treat problems of scheduling or bus operations from a traffic engineering viewpoint [65]. Unfortunately, such packages can rarely preempt buses by integrating both advanced data and system cost concepts. Snehamay Khasnabis et al have addressed a number of factors that prevent such applications in the U.S. [22]. These include the absence of a reliable technology to monitor the bus operations and to trigger preemption, lack of standards to determine warrants, and inability of the system to properly handle delay to motorists on the cross street. Therefore, the main work of this section will cover the problems of

(1) formulating a simple logic to quickly process traffic data in short time intervals.

(2) adaptively adjusting signal control timing with current real traffic information and anticipating possible impact to oncoming traffic.

4.3.1 Basic Assumptions

To begin the model development, a simplified 4-leg intersection shown in figure 4-6 is set as a basis. The east-west street which serves all bus movements and most of the total entering traffic is regarded as the major one. The south-north street is now handling passenger cars only and is regarded as the minor one. A bus route which serves two-way bus movements is settled on the east-west major street. In addition, a traffic signal used to adaptively regulate the existing traffic demand is installed at the intersection.

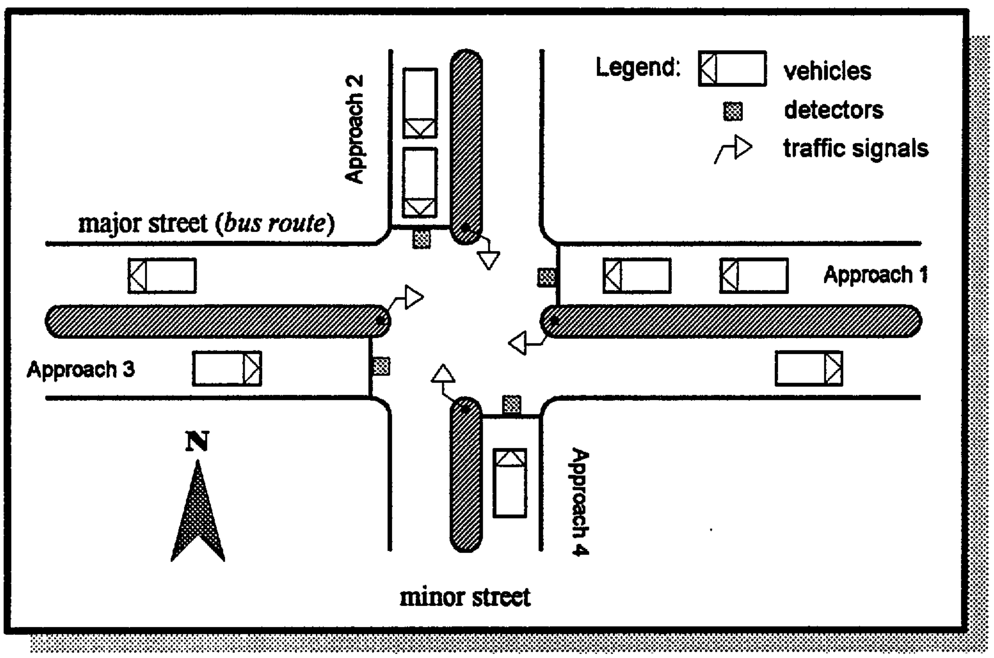

Figure 4-6 Layout of a simple signalized intersection

With such basic configurations, the following preliminary conditions presumed for characterizing the signal control model include:

1. Physical conditions:

- One-lane pavement on each of the four approaches.

- No curb parking on the intersection approaches.
- Intersection located in level terrain areas.

2. Traffic conditions:
   - Random traffic inter-arrivals on each of the four approaches.
   - Identical saturation flows or departure rates for all approaches.
   - All movements traveling straight through the intersection.
   - Queue discharge following a FCFS discipline.
   - Link travel speeds predictable at all times.
3. Signalization conditions:
   - 2-phase signal design.
   - Fully adaptive to existing traffic demand.
   - Green signal for approaches on major street adjustable and available at all time.
   - Capable of retrieving traffic information from upstream detection, communicating traffic measures (such as delay and vehicle stops), and intersection flow rates to other downstream intersections.

4.3.2 Strategy for the Signal Control Model

Although we expect to have relatively few buses compared to private cars, the buses carry relatively large number of passengers. Thus, the developed control logic is to compromise the treatments for "vehicles" and "persons" in terms of total cost. The signal system is expected to handle real time information about traffic demand and preserve road capacity for the present and/or upcoming buses concurrently.

The demand rate may actually vary randomly over time. However, it could be regarded as uniform when we emphasize some measures developed within a short time interval. With this premise, measures such as queue length accumulated at the end of each interval, total delay time or vehicular stops incurred in certain interval can be estimated step by step.

For any time interval *t*, the available traffic data may be obtained from three stages: (1) data accumulated in the past stage, (2) data obtained in the current time step, and (3) data predicted for future stage. Traffic data gathered from the past and current stages are presumed to be reliable. Also, future tragic conditions are predicted, subject to some uncertainty. Figure 4-7 shows possible traffic demand patterns in the three stages. A constant average flow rate based on historical traffic records is used throughout the entire future stage. Together with a given saturation flow rate and a preset signal timing, the demand data are used to estimate possible performance measures for the intersection in each stage.

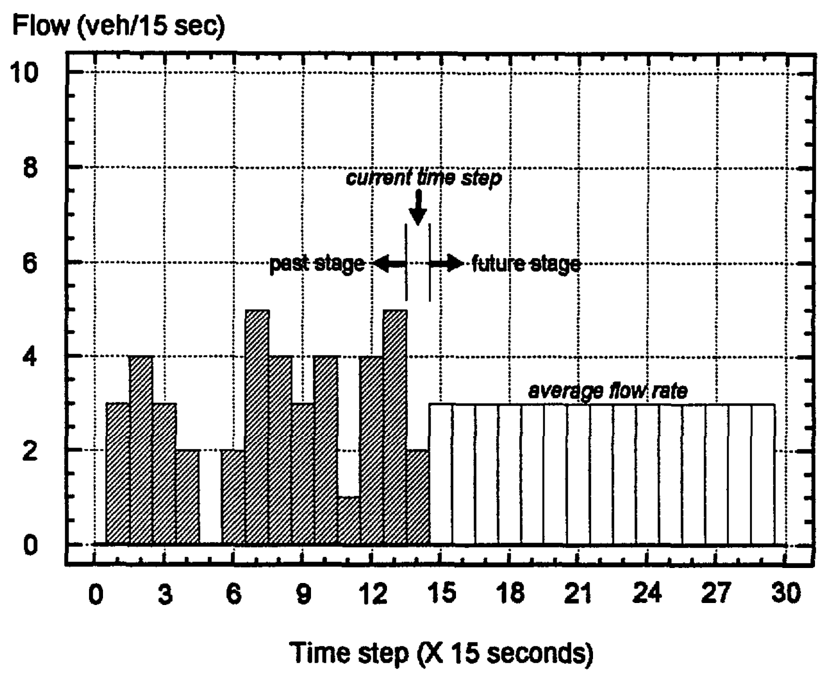

Figure 4-7 Stages for the bus priority model

Figure 4-8 demonstrates the total arrivals accumulated from each time step in figure 4-7 and a series of departures when a 60-second cycle with 50:50 green splits is applied to all stages. Traffic measures such as total delay, number of vehicular stops, and delay times w_1 and w_2 for buses 1 and 2 incurred during the future stage can be estimated from this figure.

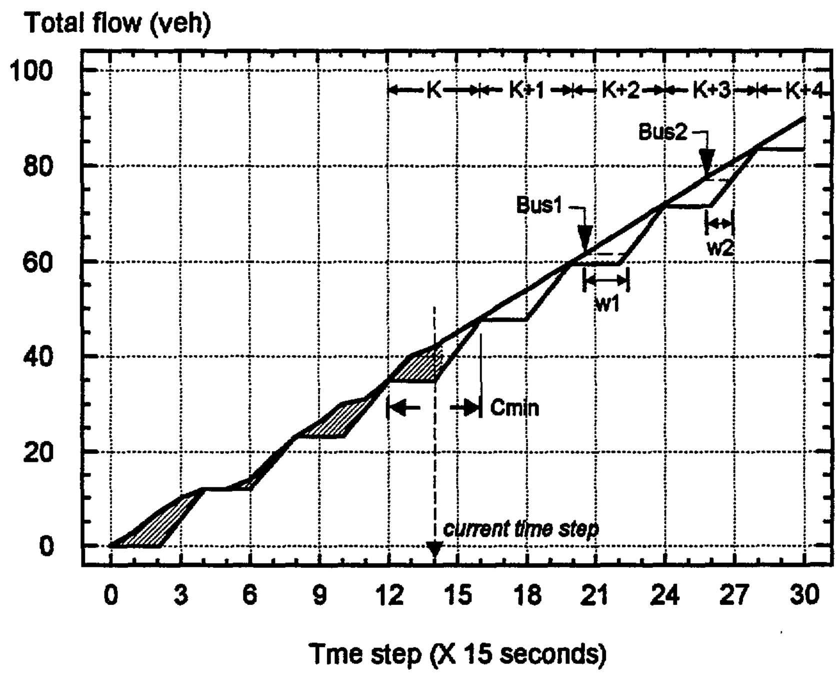

Figure 4-8 Total arrivals and discharges with fixed signal timing

It should be noted that timing design based on long term average flow rates could provide a minimum feasible cycle when X_c equals 1.0. A signal operating with minimum feasible cycle C_{min} could only provide its green phases to discharge those vehicles already in the queue. Theoretically, for a specific cycle k, if we change its duration C_{min}, all later cycles and their phasing will also be changed in order to gradually compensate for flow variations occurring in the first changed cycle k. Such transition cycles will gradually return to the initial minimum C_{min}. Figure 4-9 shows this for cases of cycle truncation and cycle expansion when a cycle k with C_{min} = 60 seconds is changed. For the cycle expansion curve, the transition cycles smoothly converge to the C_{min} right after the expanded cycle C_1' (78 seconds). For the cycle truncation curve, the cycle C_2'' that immediately follows the truncated cycle C_1'' (42 seconds) is first extended up to about 66 seconds and then the following cycles gradually approach C_{min}.

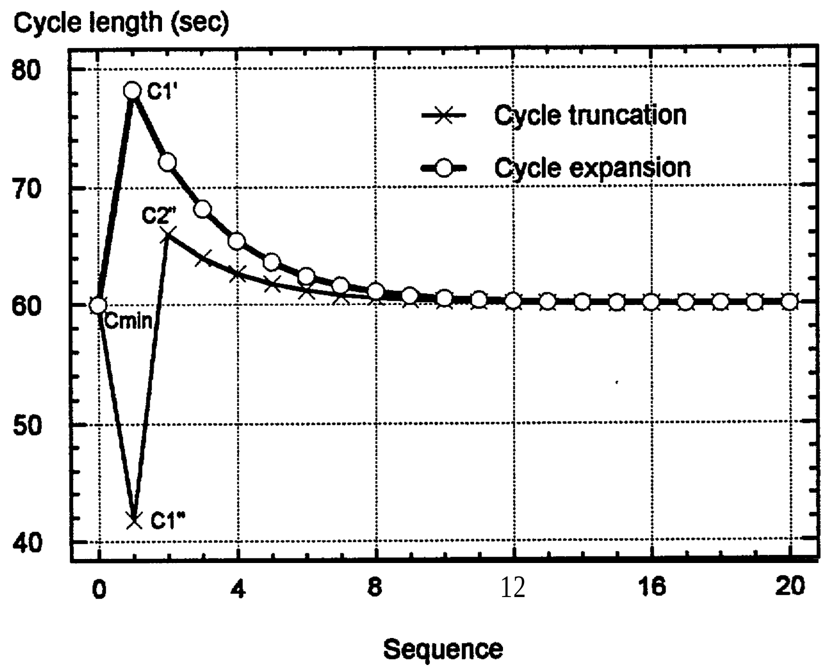

Figure 4-9 Asymptotes of cycles in transition period

Theoretically, the convergence will never end exactly at the value of initial C_{min}. The more the initial C_{min} is changed, the more the transition time is required. To overcome this problem, a tolerance ε_T that designates the absolute **difference** between each transition cycle C_l ($l = k, k+1, \ldots\ldots\ldots \infty$) and the initial C_{min} is used in this study. The convergence procedure is terminated as the ε_T is less than a pre-determined value (e.g. 0.5 second). The elapsed time required to complete a convergence procedure is called a "transition period". Thus, a transition period may contain one or more transition cycles depending on how the initial timing is changed.

To describe the possible effects of a changed timing, we extend the duration of cycle C_k (in figure 4-8) by C' seconds. This extends the duration of cycle C_k from C_{min} to $C_e = C_{min} + C'$ (see figure 4-10). The durations of all successive cycles, such as C_{k+1}, C_{k+2}, and C_{k+3}, are also changed.

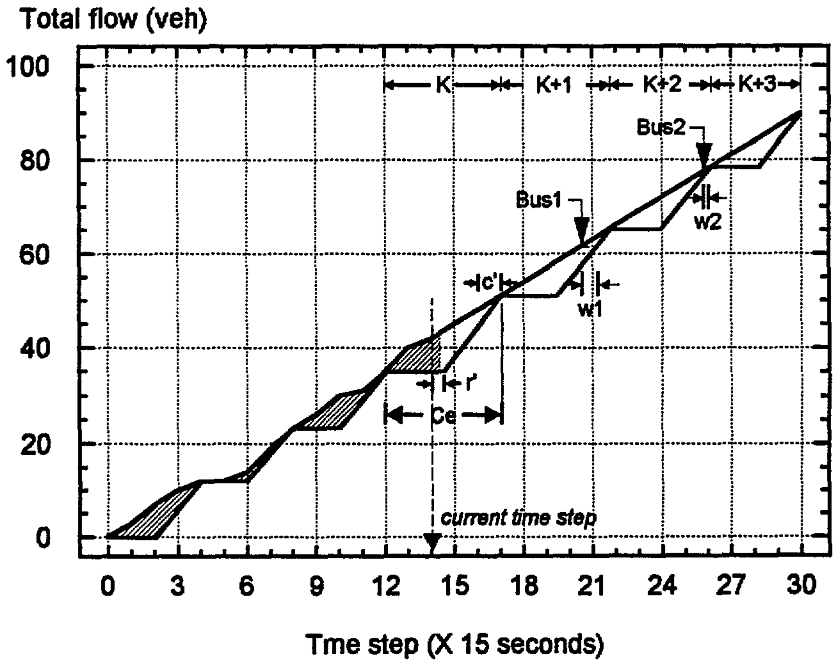

Figure 4-10 Phase adjustment for bus priority

Figure 4-10 shows two effects of these changed cycles:

(1) The expected total delay and number of stops to passenger cars during the transition period are increased. The total area of triangles involved in transition period is greater than that in figure 4-8.

(2) The expected delay to buses during the transition period is decreased. The total wait time $w_1 + w_2$ is less than that in figure 4-8.

Apparently, it is a tradeoff that any phase adjustment (either red or green) for preempting buses might incur more delays to other traffic modes. So, a control decision for adjusting current phase can be made by evaluating the TOC function. A current phase is the phase of a cycle in which the end of current time step locates. Through adjustment of current phase and its associated transition cycles, the TOC (defined as equation 4.11) of all vehicles involved in the transition period can be estimated. The signal timing with the lowest expected TOC is the one to be implemented at the next time step.

4.3.3 Effective Ranges of Phasing

The current signal phasing could be changed within limits at any time step. Theoretically, the range of phasing to be adjusted can be as wide as possible. For instance, a complete range of green time could be from g_{min} to infinite when the red time r is given. However, a longer phasing could lead to longer cycle and impose higher traffic costs compared with a shorter cycle. Moreover, it could waste most of the time searching in the wrong range. Thus, a reasonable search range of phasing (either red or green phase) is required for quickly responding to real time traffic information.

The current time step may fall in the current cycle. Thus, determination of an effective range depends on which phase the current time step is in. For a simple 2-phase signal design, the current time step t may be in either a red or a green phase. Each of the situations corresponds to its own effective range. Figure 4-11 shows the effective ranges for both situations.

1. Effective range of red phase

When the current time step t is in a red phase, two options for adjusting phasing are considered. One is the concurrent red/green phase adjustment. The other is the green phase adjustment only. A preliminary test in this study comparing the expected TOC's for both options has shown that the former imposes lower cost than the latter. Thus, the strategy would only take account of the effective range for concurrent phasing adjustment.

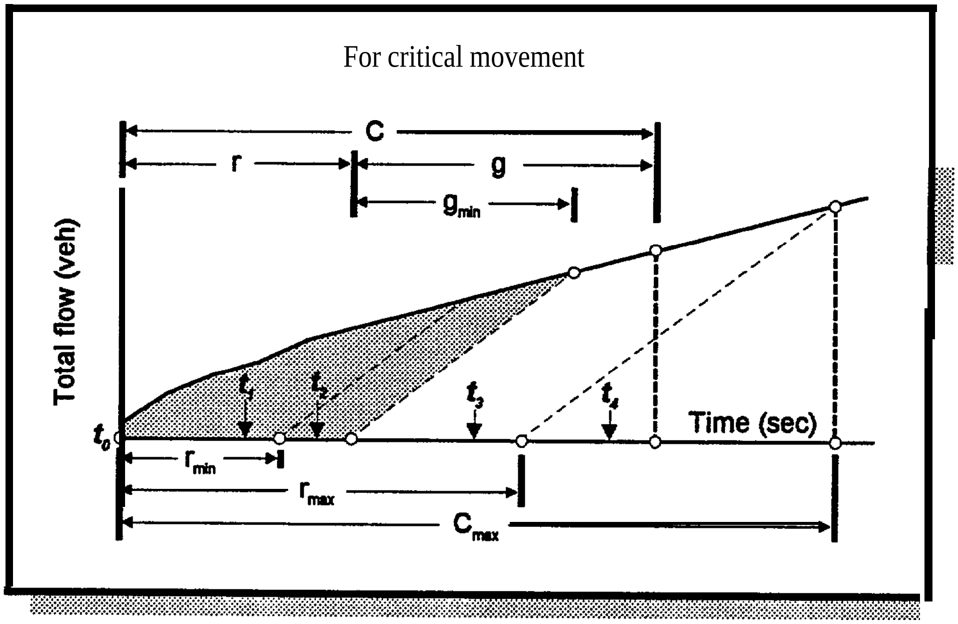

Figure 4-11 Effective ranges for phasing adjustment

The concurrent phasing adjustment means that all phases in the current cycle are changed proportionally. If the red time is determined, all the remaining phases are also determined. For a **2-phase** signal design, **from** figure 4-l 1, the effective range (a close interval) of the current red phase for one of the critical movements is

$$\Phi_R = [\mathrm{Max}\{r_{\min}, t\text{-}t_0\}, \mathrm{Max}\{r_{\min}, C_{\max}\text{-}\frac{Q_t + \overline{q}(t_0 + C_{\max} - t)}{s}\}] \quad \ldots\ldots \quad (4.12)$$

where

t Current e step

t_0 Start time of the current cycle C

Q_t Queue length at time step t, in vehicles

$\overline{q}$ Critical flow, in **veh/sec**

s Saturation flow, in **veh/sec**

$C_{\max}$ Preset maximum cycle time, in seconds

$r_{\min}$ Preset minimum red time, in seconds

The lower bound of a red phase equals r_{min} when the elapsed time from t_0 is less than r_{min} (e.g. $t = t_1$ in figure 4-11). Otherwise, the lower bound equals $t - t_0$ (e.g. $t = t_2$ in figure 4-11).

2. Effective range of green phase

When the current time step t is in a green phase, the only choice is to adjust the green phase and keep the red phase unchanged. Green phase adjustment, either extension or reduction, might temporarily cause oversaturation or impose more delay to other movements. For a **2-phase** signal design, **from** figure 4-11, the effective range (a close interval) of the current green phase for one of the critical movements is

$$\Phi_G = [\mathrm{Max}\{g_{min}\,,\ t-t_0-r\}\ ,\ \mathrm{Max}\{g_{min}\,,\ C_{max}-r\}] \quad \text{.................................} \quad (4.13)$$

where

g_{min} Preset minimum green time, in seconds

r Red time of current cycle C, in seconds

The lower bound of a green phase equals g_{min} if the elapsed time **from** t_0 is less than $r + g_{min}$ (e.g. $t = t_3$ in figure 4-l 1). Otherwise, the lower bound equals $t - t_0 - r$ (e.g. $t = t_4$ in figure 4-11).

4.3.4 Objective Function

A main goal of bus priority control is to provide transit buses with less signal delay on their routes. Generally, such goal can be attained by specially concerning the weight factors of transit buses in a cost function. The cost function, with all of its terms, is strongly related to the signal phasing. Through a cost minimization procedure, the decision variables (either red and/or green time) can be found. The newly derived variable(s) can be used as a basis to adjust the previous one(s).

The TOC function (Eq. 4.11) in section 4.2.4 accounts for a total traffic operating cost in one cycle. Such duration could be extended to a longer period containing several consecutive cycles. The transition period described in section 4.3.2 is a typical one that can be applied to the TOC function. With the same measures considered in section 4.2, the TOC function in equation 4.11 can be revised by

$$TOC_P = \sum_{l=k}^{k+n_1}[c_d(T_{d,l}) + c_s(T_{s,l})] + \sum_{l=k+n_1+1}^{k+n_1+n_2}[c_d(T_{d,l}) + c_s(T_{s,l})]$$

$$TOC_B = \sum_{j=1}^{m_1} c_{b,j}(wt_j) + \sum_{j=m_1+1}^{m} c_{b,j}(wt_j)$$

$$\text{TOC} = \frac{1}{\sum_{l=k}^{k+n} C_l}(TOC_P + TOC_B) \quad \text{.................... (4.14)}$$

where

TOC	Traffic operating cost per second, in \$/second
C_l	Duration of cycle l, in seconds
$T_{d,l}$	Total passenger car delay in cycle l, in vehicle-seconds
$T_{s,l}$	Total number of stops in cycle l, in vehicle-stops
wt_j	Expected wait time for bus j to be served in period $\sum_{l=k}^{k+n} C_l$, in bus-seconds
k	Indication of current cycle
m	Total count of buses waiting for service at time step t
m_1	Expected number of buses to be served in the transition period, $m_1 \le m$
n	Expected number of successive cycles needed to dispatch the m buses
n_1	Total number of successive cycles in the transition period, $n_1 \le n$
n_2	Total number of successive non-transitional cycles following the transition period, $n = n_1 + n_2$

c_d Unit cost for each passenger car delay, in \$/veh/second

c_s Unit cost for each vehicle stop, in \$/veh/stop

$c_{b,j}$ Unit delay cost for bus j, in \$/bus/second

4.3.5 TOC Minimization Procedure

With a set of phase boundary (range), the TOC function can be minimized to obtain the optimal control variable(s). For efficiently deriving the optimal phase, a line search procedure called "Fibonacci search" is applied in this study [66, 67, 68]. The procedure is based on the Fibonacci sequence $\{F_v\}$ defined as $F_{v+1} = F_v + F_{v-1}$ (where v = 1, 2, 3, ... and $F_0 = F_1 = 1$). The sequence is therefore 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, ... Also, the range of phasing is defined by an initial interval of uncertainty $[a_0, b_0]$. If the interval of uncertainty is $[a_k, b_k]$ at iteration k, two points of the signal phasing μ_k and λ_k can be given by

$$\mu_k = a_k + \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k) \qquad k = 1, 2, \ldots\ldots, n-1 \qquad (4.15)$$

$$\lambda_k = a_k + \frac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k) \qquad k = 1, 2, \ldots\ldots, n-1 \qquad (4.16)$$

The new interval of uncertainty $[a_{k+1}, b_{k+1}]$ is given by $[\lambda_k, b_k]$ *if* $\theta(\lambda_k) > \theta(\mu_k)$, and by $[a_k, \mu_k]$ *if* $\theta(\mu_k) \le \theta(\lambda_k)$, where θ is the TOC function to be minimized. Thus, the interval of uncertainty is reduced by the factor F_{n-k} / F_{n-k+1}. $\lambda_{k+1} = \mu_k$ if $\theta(\lambda_k) > \theta(\mu_k)$, and $\mu_{k+1} = \lambda_k$ if $\theta(\mu_k) \le \theta(\lambda_k)$.

The Fibonacci search algorithm is very similar to another one called the "Golden Section Method" (GSM). Both methods are using the concept of unimodality of function θ to reduce the interval of uncertainty [27]. The Fibonacci search differs from the GSM in that the reduction of the interval of uncertainty varies from one iteration to another. The GSM sets its new range of interval of uncertainty $b_{k+1} - a_{k+1} = \alpha(b_k - a_k)$, where the reduction ratio α is a constant 0.618. For sufficient observations n (i.e., n-m), $1/F_n$ is asymptotic to $(0.618)^{n-1}$, so that both methods are almost identical. However, when n is so limited, the Fibonacci search is more efficient than the GMS.

4.4 Cases with Signal Control for Buses

The control strategy previously developed will be tested and discussed in this section. For systematically conducting the case analyses, all of the geometric, traffic, and control conditions are consistent with those assumptions in section 4.3.

4.4.1 Signal Control for Two-way Bus Route

A two-way bus route system allows buses to travel in both directions of the route. A pre-determined headway could be maintained by these buses in either direction as they were dispatched from the upstream bus stops. Then, some of the buses might bunch up, while some would show up alone at downstream signals. Such situations limit the phasing adjustment at any time step to favor all of the concurrent bus arrivals. Basically, the decision of changing signal phasing is made by minimizing the expected TOC including costs of both buses and their riders. Therefore, buses with relatively large unit delay cost may have greater probabilities of finding green phases or reducing their delay times during red phases.

To compare the model with a no preemption condition, 250 sample buses with 5, 10, and 15 minute headways in two directions are simulated separately. A complete simulation contains 10 replication runs for each bus headway. Each replication run is terminated when a total of 250 buses is counted. To assure the results are obtained from a stable part of the simulation run, the results for the initial 50 buses from each run are discarded. A mean approach flow of 1,000 vph is assumed for q_1 and q_3 on major street (bus route) and 900 vph is for q_2 and q_4 on minor street. The population of bus inter-arrival times is randomly distributed using the given mean headway, while the unit delay cost for bus j, $c_{b,j}$, is independent and identically distributed (I.I.D.) in the interval (0.0 , 100.0]. A 60-

second cycle with volume-weighted 2-phase splits is used as the initial signal timing. Statistics of the "total bus delay" and the "traffic operating cost (TOC)" are explored in the comparisons.

The resulting total delay of buses without priority is higher than that through the bus priority model. Figure 4-12 shows the result for both models with different service headways.

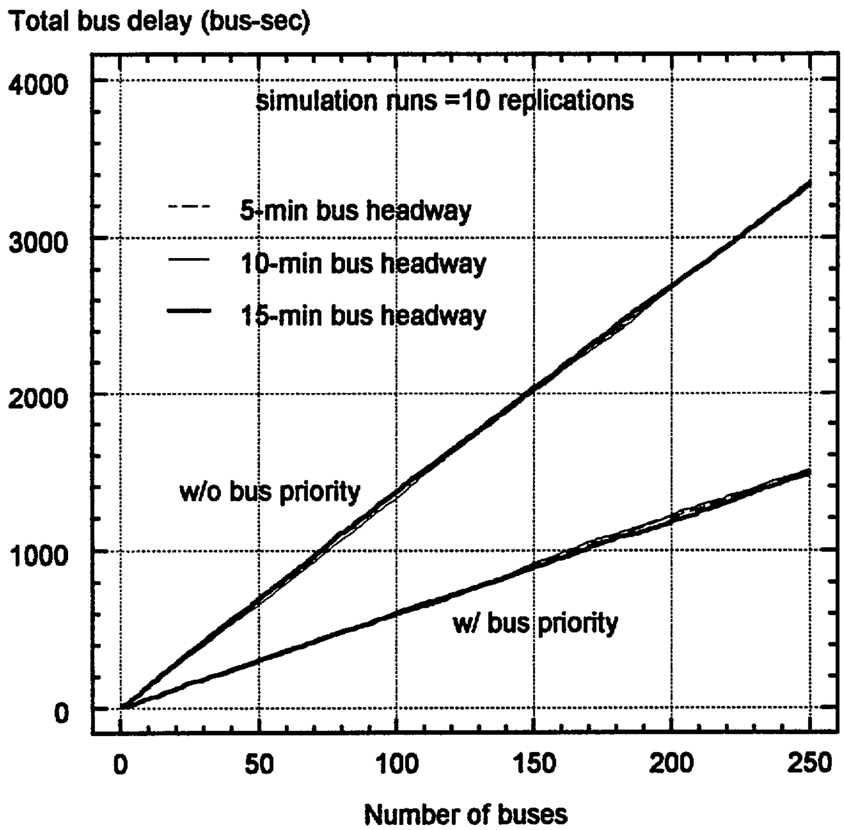

Figure 4-12 Plot of total bus delays

There are two major curve groups in figure 4-12. Each of the groups indicates that curves of bus delay for all service headways would coincide when using the same control model. The result also implies that, in the long run, bus service headways may have no effect on the improvement of total bus delay at signals. However, the bus priority model does greatly reduce total bus delay. Table 4-1 lists the associated bus delays for both controls with different bus headways. When compared with the no priority model, the improved percentage of bus delays with the priority model for any service headway can be up to 55%. This improvement rate can be referred as a basis if a signalized bus priority control is proposed by the traffic authorities.

Bus movements treated with the priority model can result in a trade-off described in section 4.3.2. It means that such a treatment will impose excessive costs in delays and number of stops to other traffic modes. Table 4-2 shows the expected TOC's for both control models with 5, 10, and 15 minute bus headways.

Table 4-1 Comparison of bus delays on the route ♦

Bus headway	Control Types	Average delay (seconds/bus)	Total delay (bus-seconds)
5 min.	w/o priority	13.39	3,346.82
	w/ priority	6.02 (-55.07%)	1,503.81
10 min.	w/o priority	13.45	3,361.86
	w/ priority	6.03 (-55.13%)	1,508.38
15 min.	w/o priority	13.36	3,338.79
	w/ priority	5.94 (-55.53%)	1,484.61

♦ Number in parenthesis denotes % change compared with the no priority model

Clearly, the bus priority model can reduce cost because it optimize the TOC function by self-adjusting the signal phasing. As the bus service headway increases, the TOC's for both controls would decrease. Shorter headways, such as 5 minutes, increase bus passages through intersections and thus require more phasing adjustments. This might break down the existing tragic condition and increase costs to all drivers and passengers at signals. Longer headways, such as 15 minutes, cause less disturbance to traffic and thus can incur lower costs compared with the shorter headways. Also, the percentage of "cost saving" decreases when bus headways increase.

Table 4-2 Comparison of traffic operating costs (TOC's) ♦

Bus headway	Control type	TOC ($/second)	Total cost ♦ (× $1,000)
5 min.	w/o priority	16.94	1,272.191
	w/ priority	15.95 (-5.84%)	1,194.636
10 min.	w/o priority	15.84	2,378.097
	w/ priority	15.32 (-3.28%)	2,296.538
15 min.	w/o priority	15.44	3,472.918
	w/ priority	15.12 (-2.07%)	3,401.930

♦ Number in parenthesis denotes % change compared with the no priority model

♦ Total cost accounts for the cost incurred during the entire simulation

4.4.2 Analysis of Bus Headways

Bus headways can affect the TOC when a bus priority control is used at signals. Frequent bus service (small bus headways) could lead to high probability of bus platooning when these buses reach the signals. On a two-way route, buses operating with long headways in opposite directions could still have a small chance of meeting at a signal in a very short time interval. For such situations, the bus priority model should be able to promptly adjust the current phasing based on the minimum TOC value.

A simulation with its initial conditions similar to those in section 4.4.1, such as 60 second cycle and the independent and identically distributed bus delay cost, is conducted by varying headways. The TOC's associated with a given bus headway are explored for both control models. Figure 4-13 shows the simulated TOC curves.

Based on figure 4-13, bus service with shorter headways will frequently interfere the current timing and incur higher TOC. However, the cost curves for both models are converging as the bus headway exceeds 40 minutes. Both cost curves asymptotically decrease to their extreme values as bus headways approach infinity. Theoretically, the extreme values for both curves should be identical and only occur when no bus disturbs the signals.

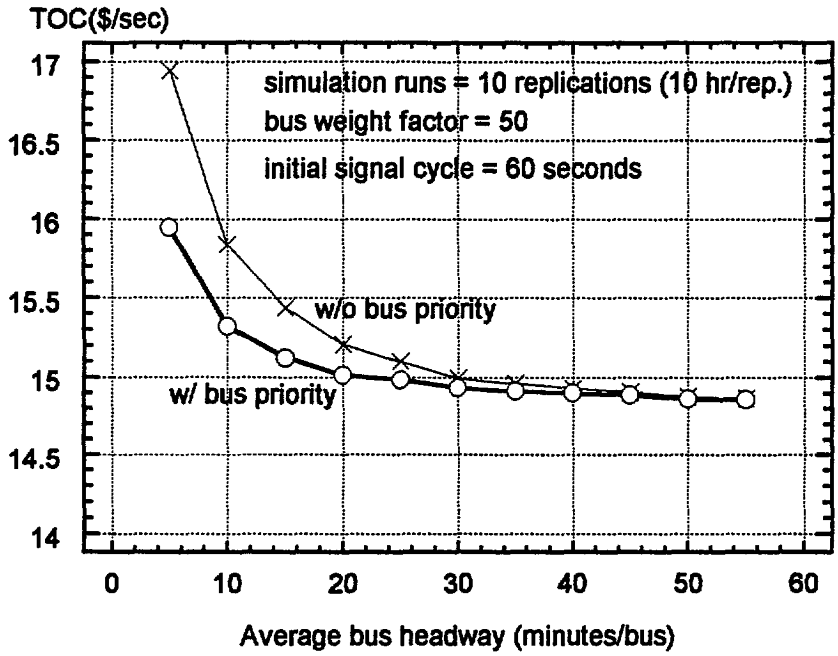

Figure 4-13 Traffic operating cost vs. bus headway

Figure 4-13 also indicates that the bus priority model for short-headway (e.g. 5-min or 10-min) bus operations reduces TOC more than the no priority model. The finding leads to the conclusion that a signalized bus priority control is especially preferred for short-headway bus service. Conversely, the effect of TOC reduction from a bus priority model can be very limited as the bus headways approach higher values (e.g. 40-min or 50-min).

4.4.3 Analysis of Bus Delay Costs

The passengers either on board or at bus stops may have different purposes for their trips and thus have different urgency levels in getting their destinations. The average time values for passengers riding buses are somewhat lower than for those in passenger cars.

However, a bus may carry ten to thirty times more passengers than a car during peak periods. Any signal-related delay to those high load buses could directly increase the TOC at the intersection. Therefore, the number of passengers on a bus can significantly affect the bus delay cost .

The time deviation from a preset bus schedule forms the other part of the bus delay cost. Due to the external interruptions or uncertainties, a bus may sometimes move ahead of or behind its expected schedule time. Intuitively, a bus which is operating far behind the schedule should obtain an immediate right of way from signals. Conversely, a bus which is operating far ahead of the schedule should be slowed down or delayed at signals, provided that effects on other traffic are considered. Therefore, the amount of time that a bus deviates from its service schedule becomes the major adjustment factor of bus wait cost.

Three cases of 5, 10, and 15 minute bus headways are simulated at varied $c_{b,j}$ values. For a given bus headway, TOC's associated with varied $c_{b,j}$ values are explored using the bus priority model. Figure 4-14 shows the three TOC curves from simulation results.

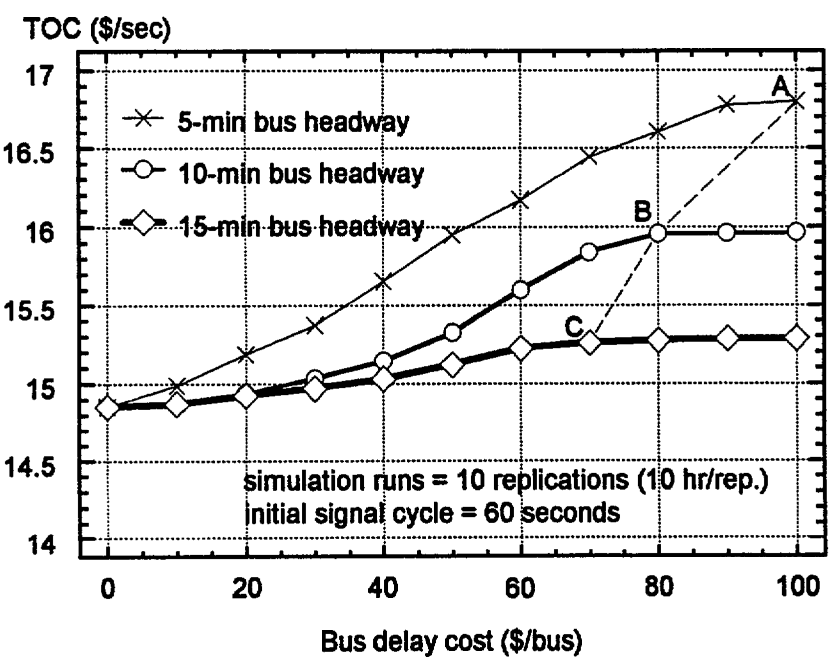

Figure 4-14 Traffic operating cost vs. bus delay cost

By inspecting a specific bus delay cost on each curve, the one with shortest bus headway incurs the highest TOC. Although the three curves are all increasing functions, they have various rates of increase. The TOC for longer headways increases slower than for shorter headways. It should be noted that the bus operating cost contributing to the TOC value decreases as the bus wait cost approaches a high value. The reason is that these highly weighted buses are always provided relatively high priority (more green or less red time) by the control model.

A phase adjustment that gives highly weighted buses with highest priority of passing will cause no delay to such buses. This means that the non-stop buses will contribute zero cost to the expected TOC value. Therefore, from each curve in figure 4-14, we can define a "stability boundary" as the highest TOC point that has the minimum $c_{b,j}$ value. A long-dash line is drawn by connecting the boundary such as point A, B, and C on each curve in figure 4-14. For a specific curve (such as the curve with 10-min headway), any $c_{b,j}$ value with its corresponding TOC point to the right side of the boundary point (such as point B) will cause no delay to the controlled buses. With this finding, the boundary line is helpful in determining which upcoming buses should be treated with absolute priority (i.e., immediate green).

4.4.4 Analysis of Signal Timing

A fixed signal with its minimal feasible cycle can minimize average delay to individual vehicles provided that the traffic demand is constant. The delay time which constitutes a major part of cost can always result a lowest TOC value when (1) the delay costs for any bus are low, or (2) no bus calls for preemption. As a signal is working with a minimum cycle setting, extending the current phasing will be the only choice for bus priority. Thus, a longer than minimum cycle could be beneficial for better phasing adjustment.

A case with an initial minimal cycle of 60 seconds (splits are determined on a volume-weighted basis) is explored with diierent bus service headways. For a given bus headway, the relations between cycle length and their corresponding TOC value are plotted in figure 4-15.

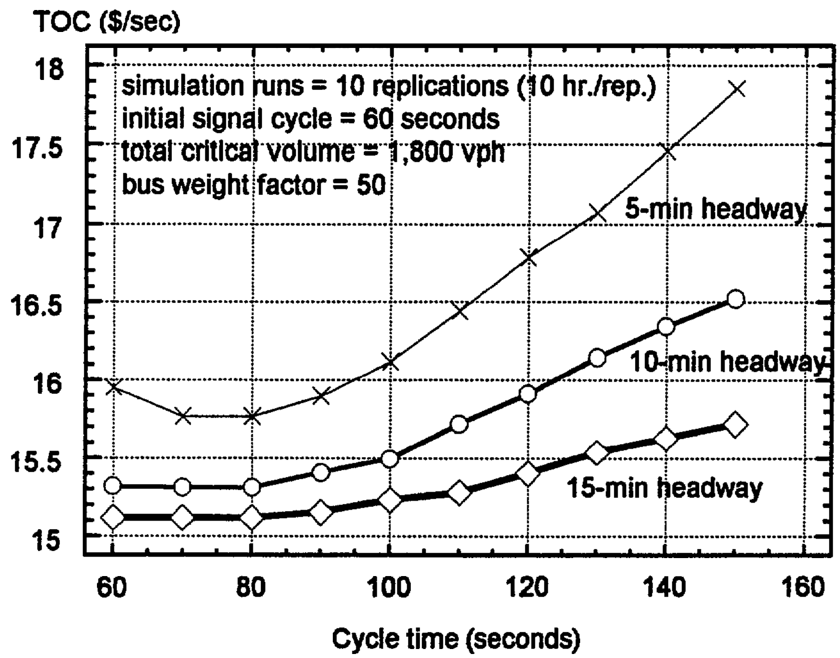

Figure 4-15 **Traffic** operating cost vs. signal cycle time

Figure 4-15 shows that, for a given cycle time, shorter **headways** incur TOC higher than longer headways. This implies that more frequent bus services (such as 5 or 10 minute headway) require more frequent signal phase changes. To obtain a relatively low TOC, longer cycles (such as 70-second cycle for both **5-min** and **10-min** bus headways) are preferable than the minimal one (i.e., **60-second** cycle). At longer **headways** (such as 15 minutes), the rarity of timing disturbances pushes the signal toward the minimal **60-**second cycle.

The trend for each curve in figure 4-15 also implies that appropriate signal timing is vital to reduce the TOC in the long run. Using a minimal feasible cycle for bus priority control may be cost-effective only when the average bus service headway is relatively long.

Chapter 5 Control of Transit Vehicles along Signalized Routes

5.1 Case Study Inputs

5.1.1 Network Configuration and Traffic Patterns

The bus dispatching control model and the signalized bus priority control model have been separately developed and tested in the previous sections. It has been shown that both control models can improve some bus service measures, such as bus travel time, dwell time, and passenger wait time, as well as optimize the control points for bus operations.

This section will test the models on a simple network which contains 22 nodes, including 12 intersections and 10 bus stops. For simplicity, each link connecting two nodes is assumed to have one lane in each direction. An east-west bus route involving all 10 bus stops and 6 of the 12 intersections is assumed to pass through the network, as shown in figure 5-l.

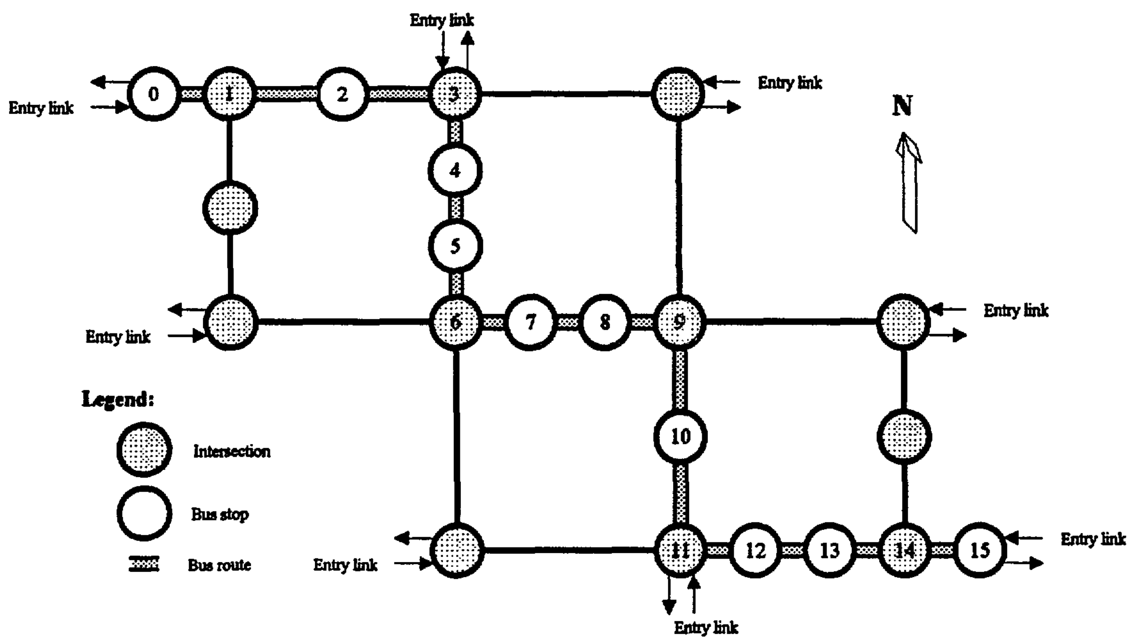

Figure 5-l Bus route on the network

With the basic configuration, a simulation model **TRAF-NETSIM** has been applied to the network. To make the simulation model work correctly, each of the bus stops along the route is coded as a dummy node from which all entering vehicles depart immediately. In addition, several initial **traffic** and control conditions, such as the input hourly volume of 500 vph on each entry link, the proportion of turning movement Rt : Thr : Lt = 1 : 2 : 1 for each intersection, and absence of signal control at any intersection, are specified before simulation begins.

The **TRAF-NETSIM** simulation model can generate intermediate **traffic** flow data and performance measures for individual intersections by intervals. The interval-based approach flow data extracted during simulation will be applied to test the bus operation control models.

A **5-minute** base interval is used to construct the approach flow table. All flow rates are expressed in terms of vehicles per hour (vph) in table **5-1**. When executing the simulation, the approach flows toward a specific intersection can be determined by referring the corresponding intersection number, approach number, and time interval in the flow table. Therefore, the approach flow rates for any intersection can be changed over time.

5.1.2 Validity of Traffic Measures

Based on the previous time-dependent flow table, two measures are collected from testing the model without any bus priority control. These two measures are total intersection delay and expected bus delay.

The average hourly approach volumes for each intersection were generated from three hour simulation runs using the **TRAF-NETSIM** model. Based on such approach volumes and a preset 20-second minimum green time, the initial fixed timing plan can be designed for each intersection. Table 5-2 lists the average hourly volumes and the 2-phase signal timings for all 6 signals along the bus route.

Table 5-1 Five-minute approach flow table

Junction 1	Minute									
Approach	**00-05**	**05-10**	**10-15**	**15-20**	**20-25**	**25-30**	**30-35**	**35-40**	**40-45**	....
1	804	792	792	804	780	792	816	828	792	....
2	-	-	-	-	-	-	-	-	-	....
3	756	912	720	816	840	912	684	900	912	....
4	336	408	432	456	480	408	492	420	408	....
Junction 3	*Minute*									
Approach	**00-05**	05-10	**10-15**	**15-20**	20-25	25-30	30-35	**35-40**	**40-45**	....
1	780	912	756	792	972	912	768	744	912	....
2	816	792	792	792	804	792	792	804	792	....
4	864	948	900	948	1056	948	1032	1032	948	....
4	720	456	504	600	480	456	480	600	456	....
Junction 6	Minute									
Approach	**00-05**	05-10	**10-15**	15-20	20-25	25-30	30-35	**35-40**	**40-45**	....
1	756	708	684	852	756	708	732	924	708	....
2	804	792	972	912	816	792	708	792	792	....
3	720	804	768	768	672	804	828	708	804	....
4	504	372	444	516	564	372	456	444	372	....
....	....									

Table 5-2 Hourly volumes and timing plans for intersections

Approach volume (vph)						
Intersection	**01**	**03**	**06**	**09**	**11**	**14**
Approach 1	**797.8**	**732.8**	**828.0**	**721.2**	**1007.4**	**908.0**
Approach 2	**-**	**800.0**	**792.4**	**448.2**	**514.2**	**450.0**
Approach 3	**872.8**	**989.4**	**746.8**	**820.2**	**786.2**	**804.2**
Approach 4	**429.8**	**538.8**	**447.2**	**835.0**	**803.8**	**-**
Signal timing (seconds)						
Cycle time	**66.61**	**103.07**	**46.90**	**46.57**	**128.38**	**66.36**
Green time 1	**40.41**	**53.67**	**20.90**	**20.10**	**68.07**	**40.36**
Green time 2	**20.00**	**43.40**	**20.00**	**20.47**	**54.31**	**20.00**

The two measures, total intersection delay and expected bus wait time, are obtained by simulating 220 buses after discarding the first 50 buses in each direction with 5-minute headways. The simulation results are compared with the theoretically expected values from fixed signal timing controls. Table 5-3 lists the comparison of total intersection delays of both theoretical and simulated results by intersections. The simulation results differ by an average of 7.68% from the theoretical ones based on uniform traffic flows.

Figure 5-2 and 5-3 show the total simulated delays in each 15-second time interval for intersections 3 and 11.

Table 5-3 Comparison of total intersection delays

Total delay (veh-sec/sec)	Intersection					
	01	03	06	09	11	14
Theoretical	6.53	20.18	9.26	9.35	25.17	6.87
Simulated	6.23	19.99	10.23	11.16	24.33	6.30
Error (%)	-4.59	-0.009	+10.48	+19.36	-3.34	-8.30

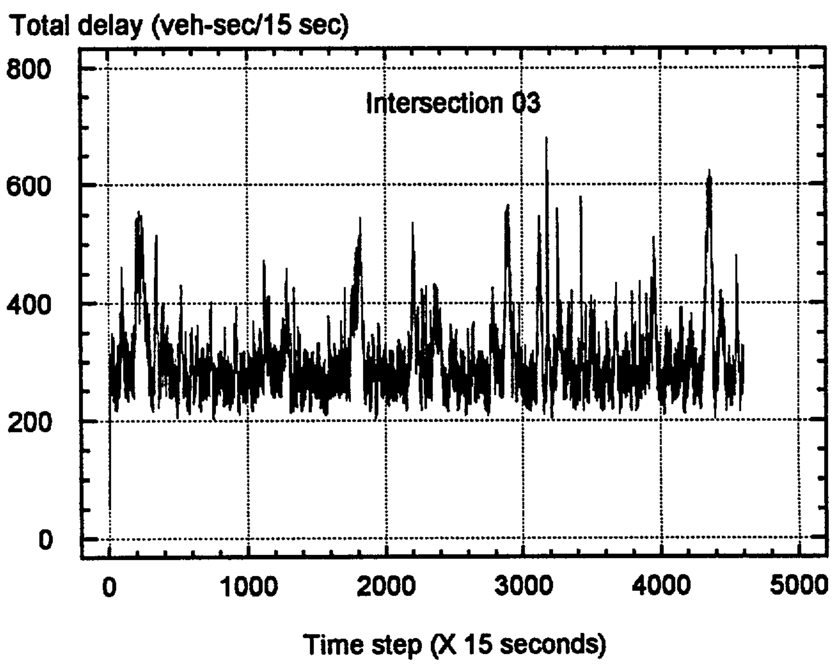

Figure 5-2 Total delay for intersection 3

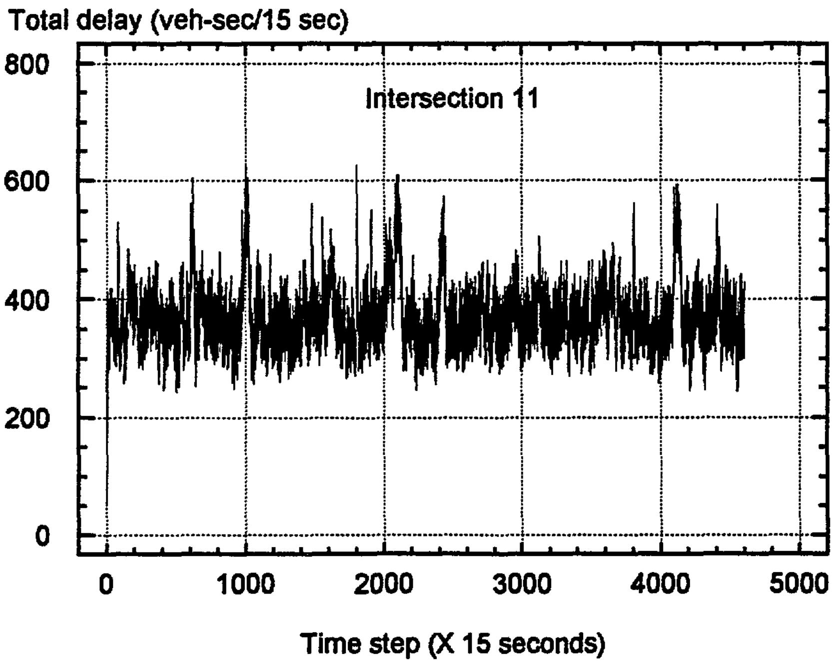

Figure 5-3 Total delay for intersection 11

The theoretical bus delay can be computed with Eq. 4.10 based on the existing signal timing and average approach volume. However, the variation of traffic demand for any intersection can not be neglected in a simulation model. The delays for some buses may increase as a temporary over-saturation or a long waiting queue occurs. Table 5-4 lists the simulated bus delays with 5-min service headways and the theoretical delays at each intersection on the east bound route. Table 5-5 lists these corresponding delays on the west bound route.

Table 5-4 Comparison of average bus delays (east bound)

Delay (bus-sec)	Signalized intersection					
	01	03	06	09	11	14
Theoretical	8.44	18.68	12.78	11.76	28.76	9.33
Simulated	8.94	19.13	23.68	14.85	30.48	9.80

Table 5-5 Comparison of average bus delays (west bound)

Delay (bus-sec)	Signalized intersection					
	01	03	06	09	11	14
Theoretical	9.00	23.64	11.50	12.56	23.34	8.52
Simulated	9.60	29.51	14.12	20.48	23.51	8.76

It should be noted that each of the simulated delays is greater than the corresponding theoretical value in both tables. This is expectable because the theoretical wait times are derived assuming no oversaturation. Yet, in the simulation, some buses may face over-saturated conditions at signals and thus may incur greater delays before they get through the signals. With such findings, the signal control model is found reasonable and is ready for the subsequent tests in this chapter.

5-2 Dispatching Controls at Bus Stops

In this section, the assumed bus route is operated with the fixed timing signal control. Six cases of headways are analyzed to investigate the effect of headways on bus movements. The same bus occupancy (i.e., the same load factor policy) is specified for these cases.

No control, headway-based control, and schedule-based control are compared for diierent headways. For the schedule-based control, the bus schedules are determined with the following approach:

$$T_i = t_0 + \sum_{j=0}^{i} (t_j^m + t_j^s + t_j^d) \quad \text{(5.1)}$$

where

T_i: The bus schedule (departure time) at node i

t_0: The bus departure time at the original terminal

t_j^m: The average moving time on jth link

t_j^s: The average stop time at nodej. If nodej is a bus stop, else $t_j^s = 0$

t_j^d: The average delay time at node j. Ifnodej is an signalized intersection, otherwise $t_j^d = 0$.

The average delay at intersections is calculated based on the optimal cycle and average traffic volume. In this approach, it is assumed that over-saturation does not exist. that assumes all vehicles arriving at an intersection during current cycle will be discharged within this green period. However, the approach can not deal with over-saturation. In fact, a bus or an other vehicle may join the queue and can not leave the intersection during this cycle due to excessive queue length. Therefore, the bus schedules determined with this approach may be faster or slower than the optimal schedules. This may affect the effectiveness of schedule-based controls.

5-2-l Regularity of Bus Operations

In our model, we assume that passenger arrivals follow a Poisson Distribution. For the Poisson distribution, the mean arrival rate of passengers is λ, and the variance is λ^2. Since we specify that load factors are equal for different headways, the smaller headway cases should have larger mean arrival rate and variation. Therefore, the number of passengers loaded by each bus tends to be equal at larger headways, so that the bus movements may have higher regularity at larger headways than at smaller headways.

The coefficient of variation is used to reflect the regularity of bus movements for diierent headway cases. The coefficient of variation for headways is defined as:

$$C_H = \frac{s}{H} \qquad (5.2)$$

where

C_H : Headway coefficient of variation

s: Headway standard deviation

H: Headway

Figure 5-4 (for east bound) and figure 5-5 (for west bound) plot the headway coefficient of variation against the headways. It can be seen that the coefficients of variation decrease as headways increase. This result is consistent with the theoretical analysis in chapter 3.

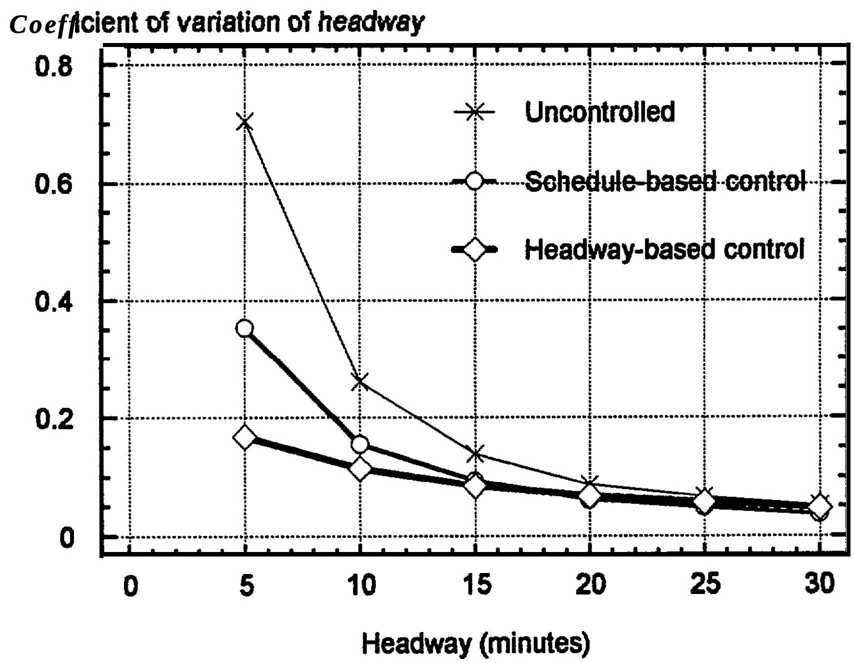

Figure 5-4 Headway variation (east bound)

It should be noted that the headway-based controls improve the regularity of bus movements more than the schedule-based controls at smaller headways. This result is consistent with the analysis in chapter 3, in which the bus route follows urban streets without signalized intersections.

In addition, at smaller headways, the implementation of bus dispatching controls has greater effectiveness. The curve with dispatching control is obviously lower than the one without any dispatching control. Nevertheless, when **headways** tend to increase, the three curves converge. This property implies that the value of controls declines at huger **headways** and the control strategies may not be necessary.

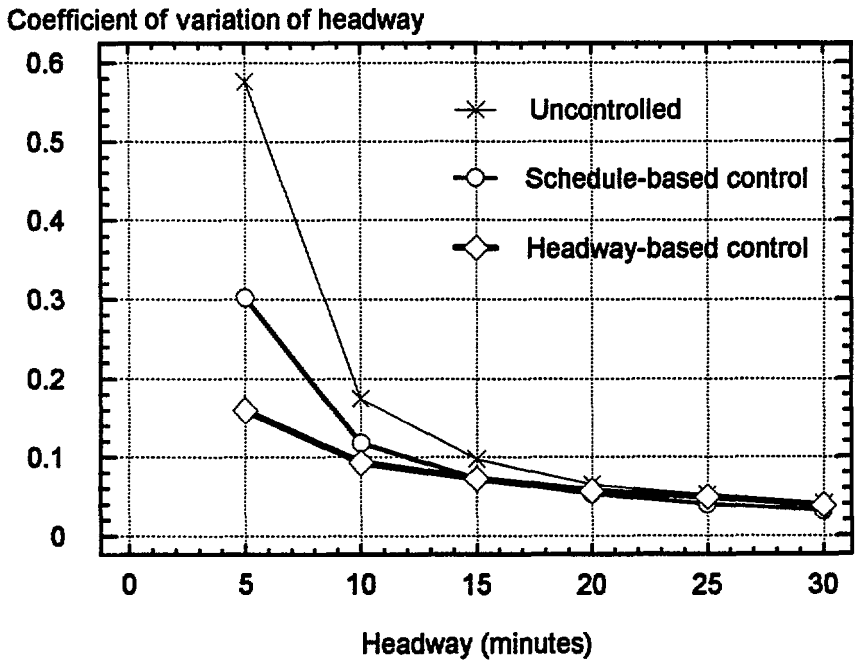

Figure 5-5 Headway variation (west bound)

5-2-2 Passenger Wait Time and Travel Time

The ideal passenger wait time curves in figure 5-6 (for east bound) and figure 5-7 (for west bound) should plot a diagonal straight line. For smaller headways, the uncontrolled curve is above the straight line, while the headway-based and schedule-based controls are approximately linear. In these two figures, the headway-based control curves are below the schedule-based control curves. This is because headway-based controls have more regular of bus movements.

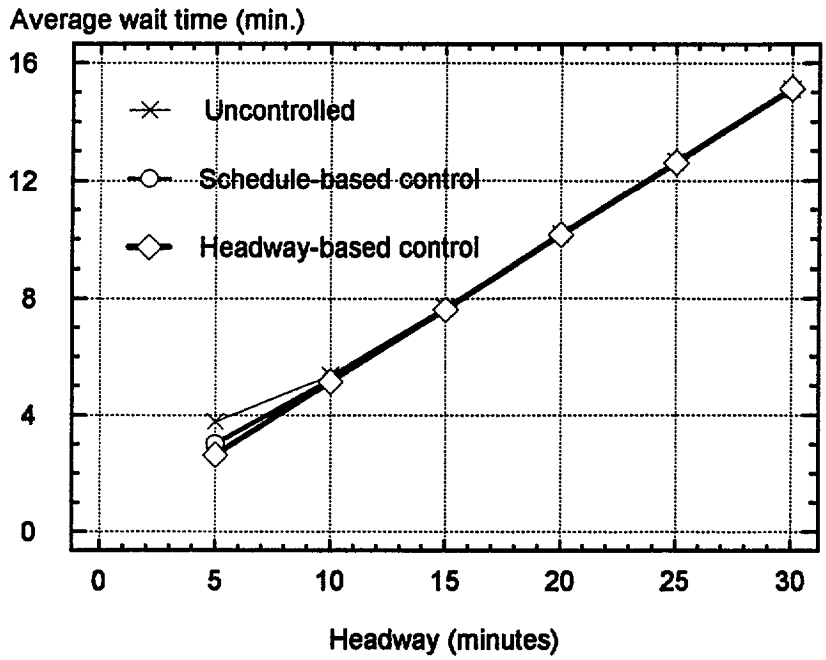

Figure 5-6 Average wait time vs. headway (east bound)

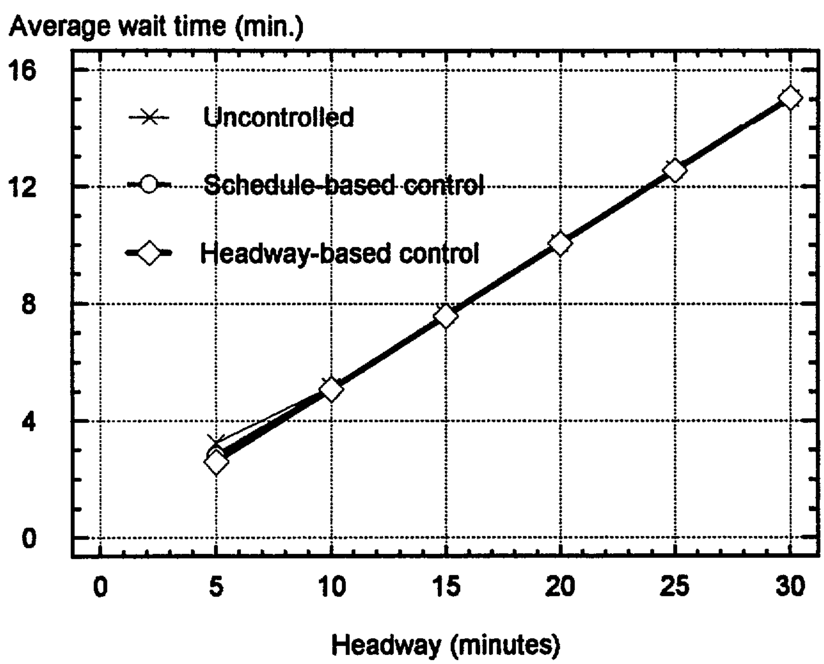

Figure 5-7 Average wait time vs. headway (west bound)

In figure 5-8 (for west bound) and figure 5-9 (for east bound), three curves for passenger travel time almost overlap, except that at smaller headway cases, the schedule-based control curve is slightly lower than the headway-based control curve, and both controlled curves are lower than the uncontrolled one.

5-2-3 Bus Travel Time

From figure 5-10 (for east bound) and figure 5-l 1 (for west bound), it can be seen that when headways are large, the bus travel times for uncontrolled, headway-based control, and schedule-based control converge. At smaller headways, the headway-based control curve indicates higher travel times than the other two curves.

5-2-4 The Effect of Bus Occupancy

In the previous analysis, the east and west bound results are presented separately. Different demands are assumed for the two directions, so that the bus occupancies diier. The load factor at the critical point is about 0.75 east bound and is about 0.5 west bound. Comparing the directional results, we can see that the east bound values (the direction with more demand) are always higher than the west bound values. This shows that larger load factors increase wait time, passenger travel time, bus travel time, and irregularity of bus operations. However, as the bus load factor approaches 1 (i.e., the bus occupancy approaches capacity), the characteristics may not hold. This is because the number of passengers loaded by each bus tends to be equal, and each bus may have an occupancy equal to bus load capacity. A more detailed analysis is provided in one of our previous studies [69].

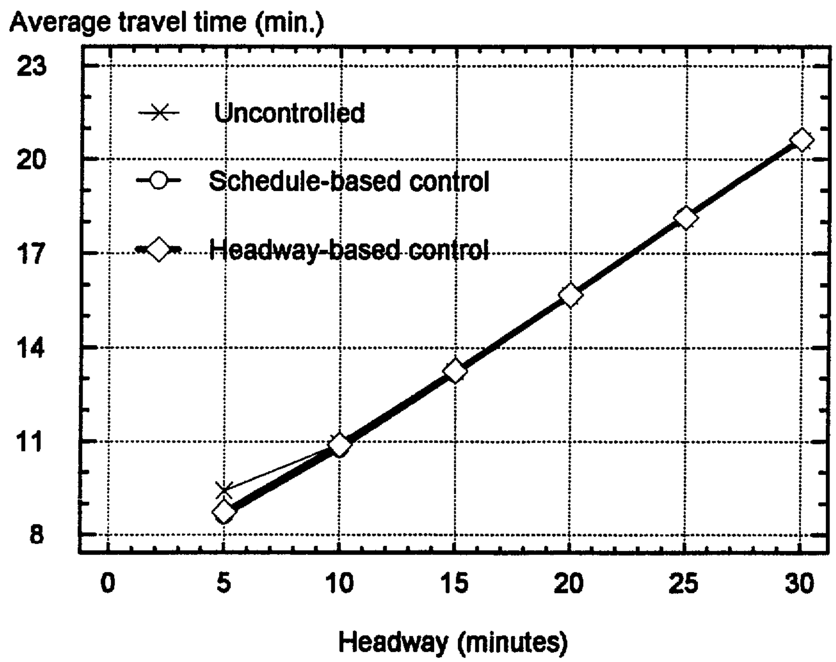

Figure 5-8 Average travel time vs. headway (east bound)

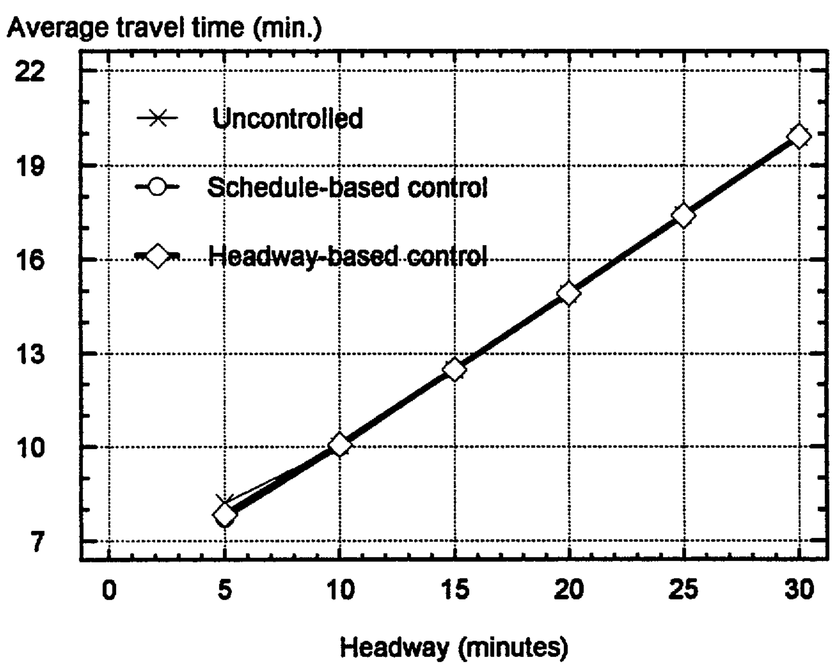

Figure 5-9 Average travel time vs. headway (west bound)

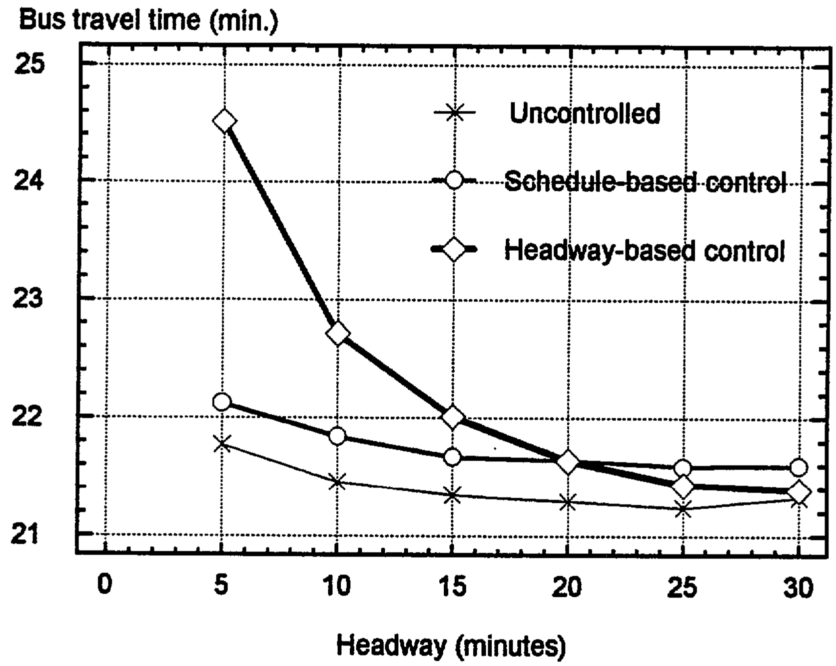

Figure 5-10 Bus travel time vs. headway (east bound)

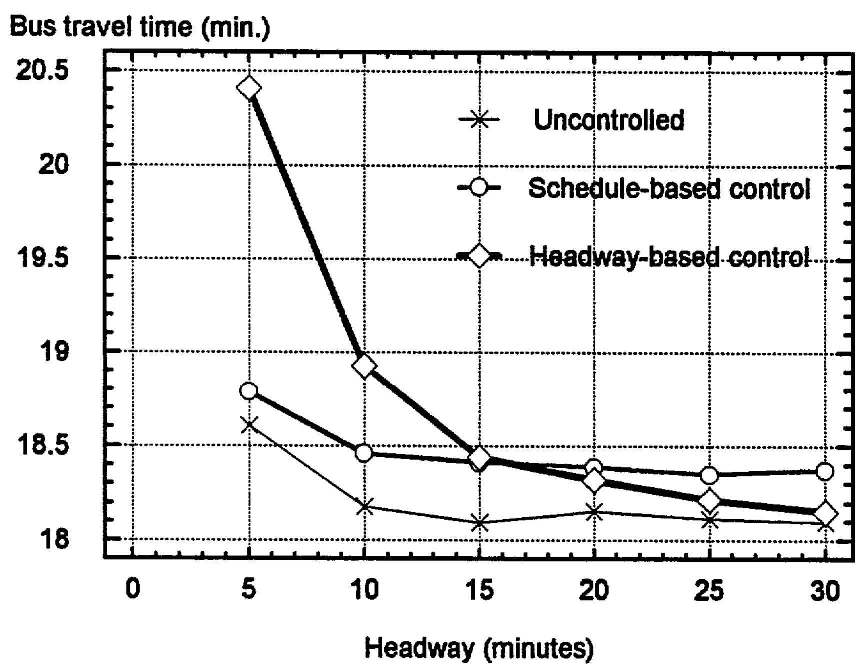

Figure 5-11 Bus travel time vs. headway (west bound)

5-2-5 The Effect of Signalized Intersections on Bus Operation

In Chapter 3, where signalized intersections are not explicitly considered, we conclude that at shorter headway cases (e.g. 5 minutes), schedule-based control strategies are preferred based on the lowest total operating cost criteria. In this section, control at signalized intersections is modeled explicitly. The stop delay at signalized intersections, effects of signals on the regularity of bus movements, passenger wait time, and passenger travel time are computed.

Tables 5-6 and 5-7 present a comparison between routes without and with signalized intersections in terms of standard deviation of headways, passenger wait time, passenger in-vehicle time, passenger travel time, and one way bus travel time.

Table 5-6 Impact of signalized intersection (east bound)

		Control strategies	Headway STD.	Wait time	In-veh. time	Travel time	Bus time
1	Signal	Uncontrolled	3.52	3.79	5.63	9.42	21.77
2		Schedule-based	1.76	3.02	5.61	8.63	22.13
3		Headway-based	0.84	2.65	6.09	8.74	24.51
4	Non-signal	Uncontrolled	3.39	3.68	5.14	8.82	20.07
5		Schedule-based	1.16	2.87	5.14	8.01	20.66
6		Headway-based	0.71	2.65	5.44	8.09	22.13
7	(1)/(4)	Uncontrolled	1.04	1.03	1.10	1.07	1.08
8	(2)/(5)	Schedule-based	1.52	1.05	1.09	1.08	1.07
9	(3)/(6)	Headway-based	1.18	1.00	1.12	1.08	1.11

From the above comparison, it can be seen that signalized intersections affect the regularity of bus movement more for schedule-based controls than for headway-based controls. For schedule-based control, the standard deviation of headway increases by about 54% due to involving of signalized intersections, while for headway-based control, by only about 20%. The passenger wait time increases by about 5% for schedule-based controls, while no change occurs for headway-based controls. However, the in-vehicle time and bus travel time increase more for headway-based controls than for schedule-based controls. This result affects on the structure of total costs.

Table 5-7 Impact of signalized intersection (west bound)

		Control strategies	Headway STD.	Wait time	In-veh. time	Travel time	Bus time
1	Signal	Uncontrolled	2.88	3.24	4.98	8.22	18.61
2		Schedule-based	1.51	2.81	4.92	7.73	18.79
3		Headway-based	0.80	2.60	5.25	7.84	20.41
4	Non-signal	Uncontrolled	2.66	3.12	4.47	7.58	16.83
5		Schedule-based	0.98	2.66	4.46	7.12	17.35
6		Headway-based	0.66	2.60	4.61	7.23	18.05
7	(1)/(4)	Uncontrolled	1.08	1.04	1.10	1.08	1.10
8	(2)/(5)	Schedule-based	1.54	1.05	1.10	1.08	1.08
9	(3)/(6)	Headway-based	1.21	1.00	1.10	1.09	1.13

5-2-6 Hourly Operating Cost

The hourly operating cost for two directions is computed based on the simulation results. Since the specified operating cost consists of passenger wait cost, passenger in-vehicle cost, and bus cost, the unit cost of the three components affects the total cost. In figure 5-12, the total cost curves are based on unit costs of $16/hour of wait time, $8/hour of in-vehicle time, and $50/hour of bus time.

The two control strategies significantly decrease the total cost at shorter headways. At longer headways, there are no significant differences between uncontrolled and the two hinds of controlled operations. In addition, the two curves of headway-based controls and schedule-based controls are very close. At smaller headways, the total cost of headway-based controls is slightly lower that of schedule-based controls. This implies that for a bus route with signalized intersections, headway-based control strategies are better than schedule-based control strategies, based on the passenger wait time and total operating cost criteria.

If we change the unit cost values of wait time from $16 to $24, the curve of headway-based controls will be lower than that of schedule-based controls at shorter headways (figure 5-1 3). This is because that the headway-based controls have the advantage of decreasing headway deviation and passenger wait times.

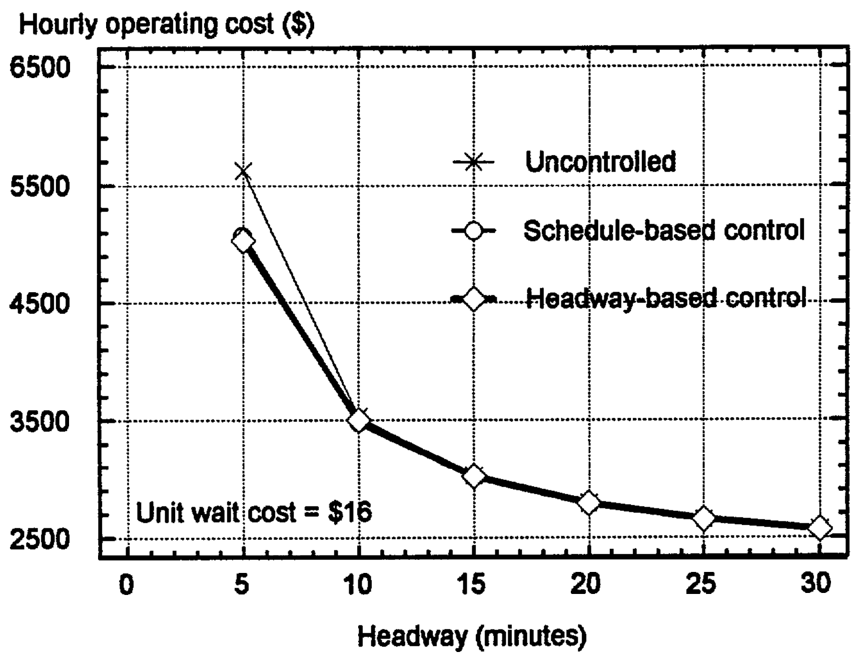

Figure 5-12 Hourly operating cost vs. headway (two directions)

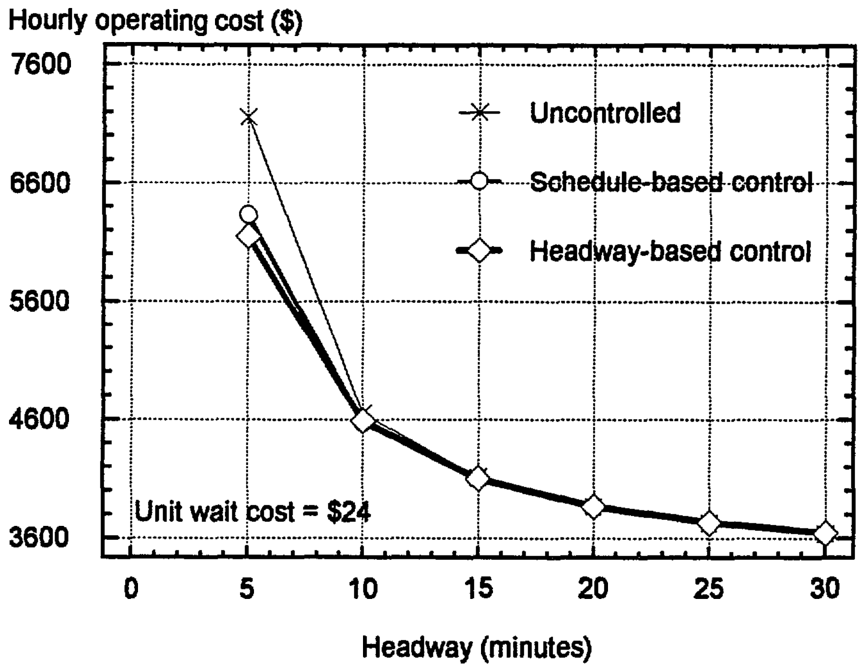

Figure 5-13 Hourly operating cost vs. headway (two directions)

Chapter 6 Adaptive Control Test Plan

6.1 Introduction

In previous chapters of this report, specific methods and algorithms for improving transit operations are presented. These algorithms use input data of various kinds to make the best transit vehicle control decisions possible. The purpose of this proposed test is to evaluate the benefits of these decisions and assess whether better ones can be made with either better input data or improved algorithms.

In general the algorithms can be tailored to accept a wide range of inputs, but the inputs available depend on the specific transit system under study. For example, a ticketing system which could provide as an input real time origin-destination data on all boarded and waiting passengers would improve dispatching decisions at transfer points relative to a conventional system which could not provide that information. In those cases where real-time data are not available the algorithms will assume values for the missing data.

In presenting this test plan the inputs available for the specific transit system used in the study are unknown so a general approach will be used. The test plan details will follow a brief summary of the basic elements of adaptive transit control, and a general procedural outline for testing.

6.2 Elements of an Adaptive Transit Control System

The transit control system given in this paper consists of inputs, computer algorithms, and outputs. Input data can take many forms; from AVL data to traffic conditions along a transit route to historical origin-destination data, etc. The input data are collected by a computer which runs algorithms to calculate the lowest cost decision(s) to make. These decisions (outputs) may be of three types: traffic signal control, transit vehicle speed control (either skipping a stop, holding at a stop, or maintaining vehicle speed between stops), or dispatching decisions for waiting transit vehicles at transfer points. As discussed in previous sections, decisions are based on minimizing a total cost function.

6.2.1 General Test Procedure

A. Select a single transit vehicle route for study.

A single route is appropriate as it is less difficult to measure than multiple routes. Preferably the route and vehicle will have some of the following features:

1. At least one transfer point along its length.
2. At least one signalized intersection along the route which allows some degree of signal timing control in real time.
3. AVL capability to allow monitoring of a transit vehicle along its route.
4. Loop detectors or other means of detecting traffic flows on approaches to signalized intersections and queue length.
5. Any additional data acquisition capabilities not mentioned above, such as passenger counters on vehicles and automatic fare boxes.
6. Any additional route-specific historical data.

B. Measure the existing route parameters without applying control (see Section 6.3, Measures of Effectiveness, for parameters to measure. Note that traffic signal related delays, items A4, A5, and A6, do not need to be measured at this step). It should be noted that this step and subsequent steps assume the use of a computer to store and time stamp incoming data (system inputs and outputs).

C. Measure route parameters (from B above) on at least one other similar, nearby route which is not expected to be directly affected by any of the control decisions made during the test. Measurements on this route will be taken during the course of the test as well. This will serve as an indicator of general transit system trends before and during the test.

D. Measure the signal control algorithm impact alone at a single signal on the route. This requires a signal with real-time controllable phases and a communication link between the bus and control center and control center and signal. For improved accuracy of the calculated cost function (and thus control decision) more information would be desirable, such as occupancy and queue detector data at all approaches to the intersection and passenger count data from the bus. The primary MOE's for comparison will be A4 and A5 from section 6.3, below.

E. Evaluate the performance of each speed control algorithm along a route.

1. Evaluate stop skipping method alone.
2. Evaluate stop holding method alone.
3. Evaluate speed control between stops method.
4. Evaluate all three of the above.

AVL data would be particularly useful to improve control decisions. Also, appropriate MOE's will be passenger wait time at stops and travel time distributions, details are specified in section 6.3.

All of the above methods require communication with the drivers, and require their adherence to the instructions. While speed control methods often are not practical given traffic and passenger demands to/from stops, if feasible, driver adherence to control input should be evaluated.

F. Evaluate the dispatching control algorithm alone at a single transfer point. For simplicity, this will preferably be a transfer with only one other route. AVL, bus occupancy data, and passenger destination data would all improve the quality of the dispatching decision to lesser or greater degrees. Transfer efficiency measures will be discussed in section 6.3.

G. Run all control methods together (4 through 7, above) and measure all route parameters.

H. Compare the results among the various trials.

6.2.2 Specific Tests

Two types of tests will be described: first stage and later stage tests. First stage tests target each control method independently (speed control, signal control, and dispatching control) to validate the control algorithms described earlier. Later stage tests combine some or all first stage tests along with refinements found while running the earlier tests.

A. First stage tests.

Test 1. Signal Control

Purpose:

To evaluate the effectiveness of a signal control algorithm which minimizes total cost for all travelers and vehicles at a single intersection.

Hardware requirements:

1. Vehicle detectors at all approaches to a selected intersection to indicate vehicle flow rates and queues by approach.
2. Hardware for the detection of bus arrival and departure to/from the selected intersection. An accurate AVL system or specially modified loop detectors that can discriminate between buses and other vehicles could be used.
3. Real-time signal timing plan modifications; local signal controllers with the ability to accept and immediately implement new signal timing plans sent from a remote controller.
4. Communication lines connecting the detectors (1 and 2) and the local signal controller to a remote controller (where timing decisions are made).
5. A data acquisition system presumably built into the remote controller to record all inputs and outputs to/from the remote controller over time. The acquisition rate should be at least 1 Hz.
6. Preferably some type of passenger counter (either APC or human counter) to indicate bus occupancy periodically by radio, although historical occupancy data, if

available, could be substituted. If radioed in, the occupancy value would need to be manually entered by the dispatcher (or other person receiving data) into a console connected to the remote controller.

Test procedure:

1. Select an intersection that meets the requirements given above.
2. Choose a two week study period during which regular traffic patterns are expected through the intersection and immediate vicinity.
3. Data acquisition system should be set to run continuously during the study period.
4. Start the test without any signal control (normal signal operation).
5. After 24 hours employ the signal control algorithm (activate the remote controller).
6. Continue to alternate every 24 hours between controlled and normal (uncontrolled) signal operation.
7. Throughout the test local traffic disturbances should be closely monitored and recorded.
8. At the end of the test period the collected data should be separated into "Controlled" and "Not Controlled" categories and statistics on the following MOE's should be manually generated (see 6.3 for calculating the values given below):

   a. Average traveler delay cost at intersection

   b. Average vehicle delay cost at intersection

   c. Average transit passenger delay cost at intersection

   d. Average non-transit passenger delay cost at intersection

   e. Total cost/hour at intersection

   f. Average transit vehicle wait time at intersection

   g. Average hourly vehicle volumes by approach

Expectations:

It is expected that the signal control algorithm will reduce average costs listed above, with the possible exception of non-transit passenger delay cost (although this is unlikely as it would indicate a very high bus to car ratio). This implies that average transit vehicle wait time is expected to decrease, and average vehicle volumes may or may not increase, but probably will increase. The above test should validate the signal control algorithm by reducing intersection costs, in particular bus and bus passenger delays. A future test would involve signal control along with speed control and/or dispatching control at transfer points to improve bus schedule adherence and transfer synchronization (to be discussed later). Those tests would use a diierent total cost function.

Test 2. Speed Control

Purpose:

This test will evaluate each of three methods of transit vehicle speed control for schedule adherence: stop holding, stop skipping, and in-route speed control.

Hardware requirements:

1. A means of recording and inputting to the control computer the location of each bus over time. This can be done in several ways:
   a. The bus driver manually radios his position to a dispatcher at every bus stop. The dispatcher immediately sends that location to the control computer, by means of a keyboard or other interface
   b. An AVL system which automatically records vehicle position over time is the better solution, if available. Some interface between the AVL hardware and the control computer would be needed for the computer to recognize the AVL data relative to the schedule (to match the AVL's (time, position) data to scheduled (time, position) and take the difference between them).
2. A means of sending speed control information to the driver.

a. The desired speed control instructions (action to take and for how long) for a given bus would be displayed on a dispatcher's monitor then relayed by radio to the driver.

b. The desired instructions would be sent directly from the control computer to a display mounted by the driver.

3. Data acquisition system presumably built into the remote controller to record all inputs and outputs to/from the control computer over time. The acquisition rate should be at least 1 Hz.

Test Procedure:

The test procedure is made up of four sub-procedures: stop holding, stop skipping, in-route speed control, all three controls simultaneously.

Procedure 1. Stop skipping:

1. Select a single bus route with the following characteristics for testing: a one way route with a minimum of 15 stops, headway of 30 minutes, and a maximum scheduled trip time of 30 minutes.
2. Choose a two week study period during which regular traffic patterns are expected along the bus route and immediate vicinity
3. Send each bus on the route at exactly the scheduled departure time.
4. Code each bus by identification number and start the data acquisition system to record each bus' movements along the route over time for the duration of the test.
5. For the first 24 hours of the test follow normal bus operations.
6. For the next 24 hours engage the control system. Continue to alternate every 24 hours between controlled and normal operation for the duration of the test.
7. Unusual traffic conditions along the route should be observed and noted by every driver at the end of each run.

8. At the end of the test period the collected data should be separated into "Controlled" and "Not Controlled" categories and statistics on the following MOE's should be manually generated (see 6.3 for calculating the values given below):

   a. Mean and standard deviation of difference between scheduled and actual arrival times at every stop.

   b. Mean and standard deviation of route travel time.

Procedure 2: Stop holding.

Procedure 3: In-route speed control.

Procedure 4: All three speed control methods simultaneously.

These procedures are essentially identical to that of Procedure 1 (above). All that differs is the type of control command sent (hold instead of skip, etc.) and the algorithms running in the control computer.

Expectations:

It is expected that stop skipping alone would reduce average travel times, stop holding alone would increase average travel times, and speed control might do either. The effect of any or all of the control methods should be to move the actual mean arrival times at stops closer to the scheduled times, and reduce the variance of arrival times. The success of some or all of these methods will indicate which methods to use for the more comprehensive later stage tests where the schedule adherence issue becomes tied directly to a total cost function.

Test 3. Dispatching control at transfer station.

Purpose:

This test will evaluate a method for improved bus transfer efficiency through real time dispatching control at a transfer point.

Hardware requirements:

The requirements here are the same as in the speed control tests (a means of tracking the buses and giving the driver timely instructions) along with other requirements.

1. A means of knowing that a transferring bus has arrived at transfer point and when. This information would probably be radioed by the driver to the dispatcher upon arrival, and the dispatcher would send the information immediately to the control computer. The control computer must also be programmed to know when transfers are scheduled to occur.
2. Preferably some type of passenger counter (either APC or human counter) to indicate bus occupancy on both the waiting and due buses. If radioed in, the occupancy values would need to be manually entered by the dispatcher (or other person receiving data) to the control computer.
3. Some prediction of transfer volumes between buses. This could come either from a modem fare collection device that knows the destinations of all passengers in the system, and relays this information to the control computer (so the computer knows how many passengers from each bus will board the other bus). Otherwise, historical transfer data, if available would need to be used. Either way, the values of expected transfer passengers must be sent to the control computer.

Test Procedure:

1. Select a single transfer point and two bus routes that connect through it with significant transfer volume and frequency (say at least 10% of total passengers arriving at transfer point transfer, and the transfer is scheduled at least once per hour during the day).
2. Choose a two week study period during which regular traffic patterns are expected along the two bus routes and immediate vicinity.
3. Code both buses by identification number and turn on the data acquisition system to record each bus' movements along the route over time for the duration of the test.

4. For the first 24 hours of the test follow normal bus operations.
5. For the next 24 hours engage the control system. Continue to alternate every 24 hours between controlled and normal operation for the duration of the test.
6. Unusual traffic conditions along the route should be observed and noted by every driver at the end of each run.
7. At the end of the test period the collected data should be separated into "Controlled" and "Not Controlled" categories and statistics on the following MOE's should be manually generated (see 6.3 for calculating the values given below):
    a. Delay cost to passengers on the ready bus (waiting).
    b. Missed connection cost to passengers on late (incoming) bus.
    c. Total cost of transfer (includes a and b above, and bus operating costs).

Expectations:

It is expected that all costs (a, b, and c) will decrease under controlled operation. If successful, this control objective can be combined with control objectives from schedule adherence and signal timing strategies to form a total cost function that includes all three objectives.

B. Later stage tests

1. Combined vehicle speed and dispatching control.

Purpose:

This test is intended to measure the combined benefits of real-time transit vehicle speed and dispatching control methods relative to normal (uncontrolled) operation for two connecting routes.

It is expected that only the most effective of the speed control methods tested in the first stage (whether stop holding, skipping, in-route speed control, or all three) will be used here.

Hardware Requirements:

These will be exactly as in the first stage dispatching control test (see Item A3 above), with the addition of the following:

Some means of counting or estimating passenger arrivals at each stop along the two routes. There are several ways to do this like relying on waiting passengers to press a button to indicate that they are waiting, or mounting video cameras to be reviewed by a person or image processor, but having human counters at each stop would probably be the most reliable and timely. Unfortunately this is an expensive option. The other possibility is to simply assume an arrival distribution based on historical or other data. This arrival information helps the computer to make the best control decision.

Test Procedure:

Although the control algorithm used is different, the procedure used will be the same as that used in the first stage dispatching control test (see Item A3 above). The MOE's will be similar, but will include additional measures. As before, the collected data should be separated into "Controlled" and "Not Controlled" categories and statistics on the following MOE's should be generated (see 6.3 for calculating the values given below):

a. Mean and standard deviation of difference between scheduled and actual arrival times at every stop (for both routes).

b. Mean and standard deviation of route travel time for each route.

c. Delay cost to passengers on the ready bus (waiting).

d. Missed connection cost to passengers on late (incoming) bus.

e. Total cost of transfer (includes c and d above, and bus operating costs).

f. Delay cost of waiting passengers at stops.

g. In-vehicle travel time cost.

h. Total system cost (includes all delay and vehicle operating costs).

Expectations:

All costs are expected to be lower under controlled operation as schedule adherence (and thus wait times) will be improved as will be transfer efficiency.

2. Combined vehicle speed, dispatching control, and signal control.

Purpose:

This test is intended to measure the combined benefits of real-time transit vehicle speed, dispatching control, and **traffic** signal control methods relative to normal (uncontrolled) operation for two connecting routes.

***Hardware** requirements:*

This will include hardware requirements for all of the first stage tests described above. In particular, the signal control hardware would preferably be incorporated into every signal along the two connecting routes; the more signals involved the lower the expected total cost to all travelers in the network.

Procedure:

Again, this test will be run with a diierent control algorithm, but will follow the procedure used in the first stage dispatching control test. As before, the collected data should be separated into "Controlled" and "Not Controlled" categories and statistics on the following MOE's should be generated. The same MOE's used above for the speed and dispatching control will be used here with the addition of delay cost to non-transit vehicles at all controlled intersections.

The above later stage tests could be directly extended to include many routes, many transfer points, and many **traffic** signals. The only difference is that there would be more routes and signals to monitor and instructions to send, while the algorithms used would be exactly the same as in the two route case.

6.3 Measures of Effectiveness

The following is a description of both the specific measures given in the test plans above and the general parameters for measurement. In general the most important measure of performance is a total cost function incorporating vehicle operating costs and delay costs to all travelers within and affected by the transit system. Almost every measure relates directly or indirectly to such a cost function.

A. Controlled-signal intersection specific measures. This category deals with measuring delay costs associated signal timing decisions. All of these values will be generated (and stored) by the control computer when calculating the signal plan to use; no calculations are required by the person running the test.

1. Average transit passenger delay cost at intersection, in $/hour. This requires knowing the time when the vehicle arrives at and clears the intersection, how many passengers are on-board, and assigning a value to the passengers' time. An AVL system or roadway detectors that indicate presence of a transit vehicle could provide the arrival and departure times at/from the intersection, as could the driver, if attentive to this task (for example push a button on a console connected by radio to the control center to indicate arrival/departure). The number of passengers aboard could be obtained from APC or bus monitors.

2. Average non-transit traveler delay cost at intersection, in $/hour. Measurement requires knowledge of the arrival and departure rates of vehicles to/from the intersection over time, average occupancy per vehicle, and the value of traveler time. While the latter two parameters can be assumed, based on historical (or other) data, the question of arrival and departure rates is more diicult. For details on the assumptions and calculations required, see the section of the report addressing the signal timing algorithm.

3. Average non-transit vehicle delay cost, in $/hour. This cost comes directly from b above, but substitute the value of vehicle time in place of traveler time.

4. Average traveler delay cost at intersection, in $/hour. This is the sum of 1 and 2 above.
5. Total system cost at intersection, in $/hour. This is the sum of 3 and 4 above and the transit vehicle operating cost (calculated as 1 above, but with value of vehicle time instead of passenger time).
6. Average transit vehicle wait time at intersection, in minutes. The wait time is already calculated in a above to get the delay cost to passengers.
7. Average hourly vehicle volume by approach, in vehicles per hour. This value comes directly from vehicle detector data; simply integrate the detector data over each hour for each approach.

B. Speed control specific measures.

1. Difference between scheduled and actual arrival times at stops, in minutes. The actual arrival times will be provided to the control computer for speed control decisions (whether by AVL or other means) and the scheduled arrival times are known. The control computer stores this difference value for each stop and route over time.
2. Difference between scheduled and actual route travel time, in minutes. This is effectively the same as 1 above, but for the final stop on the route.

C. Transfer control specific measures.

1. Delay cost to passengers on waiting bus, in $/hour. This is given by the amount of time after the scheduled departure that the bus actually leaves. This data along with the number of passengers on the bus and the value of their time gives the delay cost. The scheduled departure time is known, and the actual departure time is recorded by the control computer. The computer calculates and records this cost for every transfer.
2. Missed connection cost to late arriving bus passengers, in $/hour. As above the computer calculates and records this value. It is based on the arrival time of the

late bus, and the time until the next meeting bus. The number of passengers transferring is measured or estimated (as discussed previously) and used by the control computer for a dispatching decision.

3. Total transfer cost, $/hour. This is the sum of 1 and 2 above and bus operating costs for the waiting bus. This is calculated and recorded by the computer.

D. Other costs.

1. Delay cost to passengers waiting at stops, $/hour. The wait time of passengers could be measured directly by stationing human counters at each stop to measure arrivals over time, but more likely an average wait time will be assumed based on a predicted arrival distribution. See the previous discussion of the schedule adherence algorithm. A value of time will be assigned. The computer will record this cost.

2. In-vehicle travel time cost, $/hour. To measure this exactly would require human counters or an APC system to measure occupancy over time. Otherwise historical average occupancy data, if available, could be used. The computer will record this cost.

3. Total system cost, **$/hour**. Again this will be a sum of the above, and will be recorded by the control computer. It includes 1 and 2 above, the total transfer cost discussed in the previous section (if transfers are being performed), the total signal control cost previously discussed (if signals are being controlled in the test), and vehicle operating costs (ii not already included).

E. Miscellaneous

1. Ridership, in passengers per hour. Evaluate the mean and standard deviation of the distribution.

2. Network delays caused by disruption of a coordinated signal system (ii test is done within such a system). This may not be practical to measure.

3. Passenger perceptions. This covers such subjective issues as perceived service, security, safety, convenience, and other difficult to quantify measures. This would be measured through passenger questionnaires.
4. Fuel consumption and exhaust emissions. These are closely related costs which are not explicitly targeted by the total cost function. Although these parameters can be recorded directly for the transit vehicles, they can only be estimated for cars affected by the test.

6.4 Test Logistics

A. Coordination between transit and traffic agencies. Because this transit system test involves inputs from the roadways and control of traffic signal(s), the traffic engineering division must cooperate with the transit division to enable a successful test. The conversion of a conventional traffic signal to a real time controllable one is not a simple task and will require hardware modifications to the existing signal controller. In addition, signal wires must be run from the local signal controller to the main control center where the decision algorithms are run so that the control center can directly control the signal. This type of work is clearly within the realm of the traffic department. Furthermore the traffic department is unlikely to view the test as beneficial (particularly as any glitches in the operation of the signal will be for them to repair) and may be resistant to participating. Given this, it may help to establish an independent person or group (for example an M-IS office) to oversee the experiment with authority over both agencies on matters specific to the test.

B. Cooperation of drivers in test. It will be very important that the drivers follow the speed control decisions or the test will have little value. As mentioned before, a means of measuring driver adherence to the decisions would be helpful if possible. In general, the drivers should be exposed to the test and the importance of following the test procedure.

C. Police notification. The local police should be made aware of the test period and area. In particular they should know which signal(s) will be controlled. There should always be the ability to switch the signal(s) back to standard operation in case of any faulty operation.

D. Test Assistants. Many of the tests outlined in this plan require (or would be enhanced by) human assistance for various tasks. For example: People to count passengers on buses, and stops, people to receive incoming signals from the bus drivers in the field and translate their information into keystrokes on a keypad for the control computer to recognize.

E. Transit system coordination. Although all of the tests outlined in this plan require turning the control on and off every 24 hours, the day to day operations will be the same, and the control computer will still be turned on, receiving inputs (just not sending outputs). Therefore all that is required is to turn on/off the output from the computer as necessary, and make all drivers aware that **if they** do not receive control information, they should just run the routes normally. Otherwise, they should follow the instructions exactly.

Chapter 7 Conclusions and Recommendations

7.1 Summary

Bus priority treatments, such as provision of exclusive right of way, vehicle dispatching control, and signal preemption, have been used in various efforts to improve bus transit services. Among these, vehicle dispatching control is intended to improve passenger waiting times on bus routes. Several options, such as bus schedule adjustment, adjustment of bus headways, and turning back vehicles before the end of their route, have been found effective for route control.

Previous bus dispatching control studies provide considerable information on the analytic approaches used and actual experience obtained. Schedule adjustment studies have especially considered bus dwell time, bus running time or running speed, schedule coordination, and timed-transfer requirements. Some papers suggest that realistic schedules are very important in achieving reliable bus service. Bus headways may be adjusted by either holding early buses or skipping stops in order to either adhere to schedule or maintain more equal intervals between successive buses. Some studies have mainly focused on determining threshold values for holding and stop-skipping controls and on identifying optimal control points along routes.

Traffic signals may also be controlled to favorably influence the movements of buses. Several real-time control models with bus priority functions are reviewed in this study. These models, such as SCATS, SPPORT, and UTOPIA, treat bus movements in quite different ways. With real-time traffic information, the treatments may include green phase extension, phase early cut-off, priority-based phase allocation, or phase design with minimum traffic cost, In contrast to real-time models, an enhanced off-line model, PREEMPT, uses “need” and “eligibility” criteria to qualify a bus preemption decision. This model can be operated by using its built-in elasticity-based demand algorithm without any on-board quick-response equipment. In addition, some other simulation models such as

TRANSYT and some adaptive signal control systems, such as UTCS, SCOOT, and PRODYN are also described in chapter 2.

To improve the movements of buses along routes without incurring delays to other traffic, a new simulation model has been developed in this study. The model has the following features:

(1) Tracing the movement of each bus at all times along a bus route.

(2) Measuring the performance of bus operations in terms of passenger travel times and wait times, bus travel times, and headway regularity.

(3) Reflecting the effect of control strategies on bus performance.

(4) Considering conditions such as fluctuating traffic volumes, passenger arrivals, limited bus load capacity, and bus bunching.

(5) Estimating the costs to users as well as operators.

(6) Providing real-time information on bus movements and on-board passengers to adaptive signal control models for decision making.

With the existing signal control concepts, an adaptive signal control model is also developed by considering two basic requirements:

(1) Any bus priority decision should be based on the minimum traffic operating cost (TOC).

(2) The model should be simple enough to quickly evaluate possible traffic situations and make decisions.

The signal control model for preempting bus movements has the following features:

(1) Recording and updating traffic flow patterns occurring in the past, current and future stage in each 15 second time interval.

(2) Constructing a signal transition period based on the recorded traffic patterns and the signal timing plan.

(3) Estimating measures of effectiveness (MOE's), mainly passenger car delay, total number of stops, and bus delay.

(4) Computing and optimizing the TOC for the entire transition period.

(5) Determining the optimal signal phasing with minimum TOC for each time step.

Both the bus dispatching control and signal preemption control models are operated mainly for:

(1) Identifying the critical control variables.

For the bus dispatching control, these include the effects of holding and stop-shipping control parameters on wait time, in-vehicle time, bus travel time, and regularity of headways are analyzed and optimized based on suggested objective function. For the signal preemption control, these include the effects of adjusted phasing on the TOC function are analyzed. By using the Fibonacci search algorithm, the optimal control timing with a minimum TOC value at each time step is determined.

(2) Analyzing and comparing bus control strategies.

Headway-based control, schedule-based control, as well as uncontrolled options are compared for unsignalized and signalized bus routes, based on various criteria.

(3) Conducting sensitivity analysis.

For the bus dispatching control, the effects of headway, load factor, and time value on total cost are analyzed. For the signal preemption control, the effects of bus service headway, average time value of passengers, and signal phasing are also analyzed.

7.2 Conclusions

1. Review of system control architecture

In recent years, several methods for adjusting bus schedules or headways to achieve better route control effects have been developed. Such methods have yielded some improvements in bus travel times, dwell times and passenger wait times. Optimal control has been considered in some studies. The weaknesses found in these studies are listed as follows:

(1) Ignoring the dependent relation between bus arrivals and control strategies.

(2) Neglecting the influence of bus capacity on regularity of bus movements.

(3) Implementing controls only at limited control points, thus reducing the control effectiveness.

(4) Lacking comprehensive analyses of control effects on passenger service quality and related operating costs.

A real-time signal control model is generally considered more flexible in accommodating bus operations than a fixed-time control model. The performance measures for a bus-actuated system are better than for a fixed-time system when also considering the side street tragic. However, due to the difficulties of processing on-line data concurrently, both control models fail to treat on-line transit operations effectively. Several weaknesses have been found among the real-time control systems:

(1) Though almost all systems provide reasonable control features, some still fail to treat two or more transit vehicles coming concurrently **from** different approaches.

(2) Systems have limited capability for dynamically forecasting uncertain **traffic** patterns.

(3) Costly high-speed computers and communication systems are required for real-time signal systems.

(4) Long computation times in optimization procedures are needed to make each control decision.

(5) The reliability of O-D prediction in some models is still low for practical applications.

2. Development of bus dispatching control model

Comparisons between headway-based control and schedule-based control as well as optimization for combinations of holding and skipping controls were conducted in this study. The primary conclusions from our numerical results are as follows:

(1) The holding control parameter is the most critical decision-making variable in bus controls at bus stops. Holding control can improve significantly the regularity of bus

movement. With holding control, the average wait time of passengers decreases. However, the average in-vehicle time of passengers and the average bus travel time increases, Under schedule-based control strategies, early buses should be held until the pre-planned schedule. Under headway-based control strategies, early buses should not be held until the pre-planned headway.

(2) Stop-skipping control can be used to speed up bus movements and regulate headways. However, tight stop-skipping control significantly increase average wait time. The experimental results from several patterns of load factors and **headways** show that skipping control does not significantly decrease either user cost or total costs. However, overly tight holding control may make things worse. Therefore, stop-shipping control is not recommended.

(3) A headway-based control strategy has advantages compared to schedule-based control in improving the regularity of bus movements and reducing wait times. Its disadvantages are increases in both bus and passenger travel times. Hence, **headway-**based control must tradeoff between wait time and travel time. In addition, with headway-based control, the location and departure time of the preceding bus should be sent to the following bus. Thus, suitable communication equipment is needed on buses.

(4) Schedule-based controls have advantages compared to headway-based control in improving passenger in-vehicle times and bus travel times. Such controls are easy to implement because they do not need **information** of bus locations. In addition, schedule-based controls improve the on-time performance of bus service. This is very important for long headway situations. Thus, schedule-based control strategies are strongly recommended.

(5) Bus controls yields greater benefits at higher bus frequencies and load factors than at lower frequencies and load factors. At low loads and low frequencies, uncontrolled operation would not be worse than controlled operation.

(6) The deviation of headway (a measure of bus movement regularity) should not be taken as a unique decision-making criterion. Greater regularity of bus services may not be consistent with lower passenger cost, A narrower control range would yield more regular movement of buses, but increase passenger wait time and travel time.

(7) The total cost is a more comprehensive and hence preferable decision-making objective. It includes both user cost and supplier cost. The total cost is influenced by fraction of user cost and supplier cost, value of wait time and in-vehicle time, and passenger demand.

3. Development of signal preemption model for buses

Comparisons of operations with and without bus priority controls were made in this study. The main results are as follows:

(1) The total bus delay without bus priority is higher than with the bus priority model. For any bus headway, our priority model can significantly improve bus delays, by up to 55%.

(2) Bus priority may impose excessive operating costs, such as delay or vehicular stops, to other traffic modes. As the scheduled bus headway increases, the traffic operating cost (TOC) decreases both with and without priority controls. However, the difference between the two TOC's decreases as the bus headway increases. This implies that bus priority control is preferable for short bus headways. As the bus headways get large, the TOC saving from bus priority controls is very limited.

(3) In the long run, scheduled bus headways may have no effect on the improvement of total bus delay at signals. However, dense bus platooning or concurrent arrivals cause frequent changes in signal phasing and thus increase costs to other traffic. Thus, the TOC of bus priority control with long bus headways is lower than that with short bus headways.

(4) With a bus priority control, the bus operating cost contributing to TOC decreases as the unit bus delay cost increases. This is because the control model causes lower delay to the buses with higher unit delay costs.

(5) As shown in figure 4-14, a boundary condition can be developed to determine which buses should receive absolute priority (i.e., immediate green) and which buses should receive a lower priority treatment.

(6) To obtain a relatively low TOC, longer cycles than the minimal one are preferable. However, as the bus headways increase, the rarity of timing disturbances tends to restore minimal cycles.

(7) Appropriate signal timing can reduce the TOC in the long run. Using a minimal feasible cycle for bus priority controls might be cost-effective if the average bus headway is extremely long. Otherwise, a longer basic cycle is preferable.

7.3 Recommendations for Future Research

(1) Optimization of bus fleet size

In this model, bus layover time is a deterministic parameter used to calculate the required number of buses. It is assumed that buses can be available at the terminal anytime. In practice, layover time available to buses is a random variable due to the uncertainty of bus arrival time at terminal stations. Thus, if a bus returns late to the original terminal, it can not be dispatched on time. Therefore, the developed model should be improved to realistically determine bus fleet size under probabilistic conditions.

(2) Prediction of bus arrival times

The bus arrival time distribution is useful for dispatching buses, especially in a timed transfer transit system. A bus arrival prediction model should be developed to take advantage of traffic monitoring and automatic vehicle location systems.

(3) Prediction of passenger demand and ridership

Improved models should be developed to estimate the passengers waiting for buses, the passengers on board buses who would be delayed by control decisions and the passengers wishing to transfer to other routes. It would be especially useful to integrate our models with standard demand forecasting models.

(4) Improved control at transfer stations

Bus controls should be oriented toward maintaining the regularity of bus operation on their route and minimizing connection costs at transfer stations.

(5) Model the operation and control of light rail transit (LRT)

The models developed here could be modified relatively easily to handle LRT operations through signalized intersections and at transfer stations.

(6) Improved traffic simulation

It would be desirable to integrate the transit control models for buses and LRT with traffic simulation models to improve the prediction of traffic conditions and travel times in congested networks.

References

1. Ercolano, J. M., *'Limited-stop Bus Operation: An Evaluation"*, Transportation Research Board, Transportation Research Record 994, 1984, pp. 24-29.
2. Furth, P. G. and Day, F. B., *"Transit Routing and Scheduling Strategies for Heavy Demand Corridors"*, Transportation Research Board, Transportation Research Record 1011, 1985, pp. 23-26.
3. Turnquist, M. A., *'A Model for Investigating the Effects of Service Frequency and Reliability on Bus Passenger Waiting Time"*, Transportation Research Board, Transportation Research Record 663, 1978, pp. 70-73.
4. Transportation Research Board, *"Supervision Strategies for Improved Reliability of Bus Routes"*, National Cooperative Transit Research & Development Program, Synthesis of Transit Practice No. 15, 199 1.
5. Furth, P. G., *"Short Turning on Transit Routes"*, Transportation Research Board, Transportation Research Record 1108, 1987, pp. 42-52.
6. Jordan, W. C. and Tumquist, M. A., *"Zone Scheduling of Bus Routes to Improve Service Reliability"*, Transportation Science, Vol. 13, 1979, pp. 242-268.
7. Brand, D., Attanucci, J. and Morris, H., *"Southeast Expressway Reserved Lane for Buses and Carpools"*, Transportation Research Board, Transportation Research Record 663, 1978, pp. 76-87.
8. Organization for Economic Cooperation and Development, *"Optimization of Bus Operation in Urban Areas"*, 1972.
9. Abkowitz, M. D. and Engelstein, I., *"Temporal and Spatial Dimensions of Running Time in Transit System"*, Transportation Research Board, Transportation Research Record 877, 1982, pp. 64-67.
10. Levinson, H. S., *"Analysis Transit Travel Time Performance"*, Transportation Research Board, Transportation Research Record 915, 1983, pp. l-6.
11. Abkowitz, M. and Engelstein, I., *"Method for Maintaining Transit Service Regularity"*, Transportation Research Board, Transportation Research Record 961, 1984, pp. l-8.
12. Guenthner, R. P. and Sinha, K. C., *"Modelling Bus Delay due to Passenger Boarding and Alighting"*, Transportation Research Board, Transportation Research Record 915, 1983, pp. 7-13.
13. Osuna, E. E. and Newell, G. F., *"Control Strategies for an Idealized Public Transportation System"*, Transportation Science, Vo16, No. 1, 1972, pp. 52-72.

14. Koffman, D., *"A Simulation Study of Alternative Real-time Bus Headway Control Strategies"*, Transportation Research Board, Transportation Research Record 663, 1978, pp. 41-46.
15. Tumquist, M. A. and Blume, S. W., *"Evaluating Potential Effectiveness of Headway Control Strategies for Transit System"*, Transportation Research Board, Transportation Research Record 746, 1980, pp. 25-29.
16. Abkowitz, M., Eiger, A. and Engelstein, I., *"Optimal Control of Headway Variation on Transit Routes"*, Journal of Advanced Transportation, Vol.20, No. 1, 1986, pp. 73-88.
17. Tumquist, M. A., *"Strategies for Improving Reliability of Bus Transit Service"*, Transportation Research Board, Transportation Research Record 8 18, 1978, pp. 7-l 3.
18. Seneviratne, P. N., Tam, L. and Javid, M., *"Scheduling Fixed Route Bus Services Using Simulation"*, Proceedings of the International Conference on Microcomputers in Transportation, 1989, pp. 1042-l 053.
19. Seneviratne, P. N., *"Analysis of On-time Performance of Bus Services Using Simulation"*, Journal of Transportation Engineering, Vol. 166, No.4, 1990, pp. 5 17-531.
20. Abkowitz, M. D. and Lepofsky, M., *"Implementing Headway-based Reliability Control on Transit Routes"*, Journal of Transportation Engineering, Vol. 116, No. 1 1990, pp. 49-63.
21. Transportation Research Board, *"Bus Route and Schedule Planning Guideline"*, National Cooperative Research Highway Program, Synthesis of Highway Practice No.69, 1980.
22. Khasnabis, S., Reddy, G. V. and Hoda, S. K., *"Evaluation of the Operating Cost Consequences of Signal Preemtion as an IVHS Strategy"*, Transportation Research Record 1390, 1993.
23. Lowrie, P. R., *"The Sydney Co-ordinated Adaptive Traffic System - Principles, Methodology, Algorithms"*, Deptment of Main Roads, N.S.W., Australia, 1982.
24. Sims, A. G. and Finlay, A. B., *"SCATS, Splits and Offsets Simplified (S.O.S.)"*, ARRB Proceedings Vol. 12, Part 4, 1984.
25. Longfoot, J. E., *"SCATS - Development of Management and Operation Systems"*, International Conference on Road Traffic Signalliig, London, 1982.
26. Luk, J. Y. K., Sims, A. G. and Lowrie, P. R, *"SCATS - Application and Field Comparison with a TRANSYT Optimised Fixed Time System"*, International Conference on Road Traffic Signalling, London, 1982.

27. Cornwell, P. R., "*Dynamic Signal Co-ordination and Public Transport Priority*", Road Traffic Authority, Australia, 1986.

28. Han, B. and Yagar, S., "*Real-time Control of Traffic with Bus and Streetcar Interactions*", University of Waterloo, Canada, 1991.

29. Yagar, S., "*Efficient Transit Priority at Intersections*", Transportation Research Record 1390, 1993.

30. Bell, M. G. H., "*Future Directions in Traffic Signal Control*", Transportation Research Vol. 26A, 1992.

31. Khasnabis, S., Reddy, G. V. and Chaudry, B. B., "*Signal Preemption as a Priority Treatment Tool for Transit Demand Management*", *VNIS*, 199 1.

32. El-Reedy, T. Y. and Ashworth, R, "*The Effect of Bus Detection on the Performance of a Traffic Signal Controlled Intersection*", Transportation Research, Vol. 12, 1987.

33. Redwan, A. E. and Benevell, D. A., "*Bus Priority Strategy: Justification and Environmental Aspects*", Journal of Transportation Engineering, Vol. 109, No. 1, 1983.

34. Jacobson, J. and Sheffi, Y., "*Analytical Model of Traffic Delays under Bus Signal Preemption: Theory and Application*", Transportation Research B, Vol. 15B, 198 1.

35. Rossi, T. F. and Ceder, A., "*Signalized Intersections with Variable Flow Rates: I - Analysis and Simulation*", Journal of Advanced Transportation, Vol. 22, No.2, 1988.

36. Rossi, T. F. and Ceder, A., "*Signalized Intersections with Variable Flow Rates: II - a Model for Delay Estimation*", Journal of Advanced Transportation, Vol. 23, No.1, 1989.

37. Su, C. C., '*An Adaptive Traffic Signal Control System*", Ph. D. Dissertation, University of Maryland at College Park, 1994.

38. Henry, R. D. et *al*, "*Evaluation of UTCS Control Strategies - Exclusive Summary*", Report No. FHWA-RD-76-149, DOT, Washington, D.C., 1976.

39. Gabard, J. F. et *al*, "*CALIFE: On-line Calculation of Fixed Time Plans*", The 5th IFAC/IFIP/IFORS International Conference on Control in Transportation Systems, Vienna, Autriche, 1986.

40. Gabard, J. F., "*CALIFE: An Adaptive Strategy for the On-line Calculation of Traffic Signal Plans in Urban Networks*", Recherche Transports Securite (English Issue), 1990.

41. Hunt, P. B. and Robertson, D. I. et *al*, "*SCOOT - A Traffic Responsive Method of Coordinating Signals*", Department of Transport, TRRL Report, LR 1014, Crowthome, 198 1.

42. Hunt, P. B. and Robertson' D. I. et *al*, "*The SCOOT On-line Traffic Signal Optimization Technique*", International Conference on Road Traffic Signalling, London' 1982.
43. Clowers, D. J., "*Real Time Area Traffic Control - the User's Viewpoint of SCOOT*', International Conference on Road Traffic Signalling, IEE, London, 1982.
44. Carden, P. and McDonald, M., "*The Application of SCOOT Control to an Isolated Intersection*", Traffic Engineering and Control, Vol. 26, No. 6, 1985.
45. Bretherton, R. D. and Brown, G. T., "*Recent Enhancements to SCOOT: SCOOT Version 2.4*", IEE, Road Traffic Monitoring and Control, London' 1990.
46. Chard, B. M. and Lines, C. J., "*TRANSYT: the Latest Developments*", Traffic Engineering and Control, Vol. 28, 1987.
47. Vincent, R. A. and Young, C. P., "*Self-optimization Traffic Signal Control Using Microprocessor: the MOVA Strategy for Isolated Intersections*", Traffic Engineering and Control, Vol. 27, No. 7, 1986.
48. Vincent, R. A. and Peice, J. R., "*MOVA: Traffic Responsive, Self-optimizing Control for Isolated Intersections*", TRRL LR 170, Crowthorne, 1988.
49. Gartner, N. H., "*OPAC: A Demand Responsive Strategy for Traffic Signal Control*", Transportation Research Board, Transportation Research Record 906, 1983.
50. Henry, J. J., Farge, J. L. and Tuffal, J., "*The PRODYN Real Time Traffic Algorithm*", IFAC-IFIP-IIFORS Conference on Control in Transportation' 1983.
51. Farges, J. L., Khoudour, L. and Lesort, J. B., "*PRODYN: On Site Evaluation*", IEE Conference Publication 320, UK, 1990.
52. Mauro, V. and Taranto, D., "*UTOPIA*", IFAC-IFIP-IFORS Conference on Control, Computers, Communication in Transportation, 1989.
53. Vuchic, V. R., "*Urban Public Transportation, System and Technology*", Prentice-Hall, Inc., 1981, pp. 200-201.
54. Kraft, W. H., '*Bus Passenger Service -Time Distribution*", Transportation Research Board, Transportation Research Record 625, 1977, pp. 37-43.
55. Liang, P., "*Simulation of Bus Control Strategies*", M.S. Thesis, University of Maryland at College Park' 1994.
56. Lee, J. and Chou, Y. W., "*An Uniform and Simplified Timing Design Procedure for Pre-timed Signals*", Vol. 1, Proceedings of the 3rd Conference of REAAA, 198 1.
57. Pignataro, L. J., "*Traffic Engineering Theory and Practice*", 1973.
58. Webster, F. V. and Cobbe, B. M., "*Traffic Signals*", Road Research Laboratory, Road Research Technical Paper 56, 1976.

59. Wohl, M. and Martin' B. V., "*Traffic System Analysis for Engineers and Planners*", 1972.

60. Akcelik, R., "*Traffic Signals: Capacity and Timing Analysis*", ARRB *No.* 123, Australian Road Research Board, 198 1.

61. Hsu, T. P., "*Traffic Signal Timing Design and Analysis at Isolated Intersections*", Transportation Planning Journal, Vol. 14, No. 2, 1985.

62. Transportation Research Board, "*Highway Capacity Manual*", Transportation Research Board, TRB Special Report 209, 1985.

63. Li, J., Rouphail, *N.* M. and Akcelik, R., "*Overflow Delay Estimation for a Simple Intersection with Fully-actuated Signal Control*", TRB Paper No. 940610, 1994.

64. Akcelik, R., "*Estimation of Green Times and Cycle Time for Vehicle-actuated Signals*", TRB Paper No. 940446, 1994.

65. PC-RANS, "1994 *PCs in Transportation Software Directory*", PC-TRANS, Kansas University Transportation Center, 1994.

66. Bazaraa, M. *S.* and Shetty, *C.* M., "*Nonlinear Programming - Theory &Algorithms*", John Wiley & Sons, 1990.

67. Avriel, M., "*Nonlinear Programming, Analysis and Methods*", Prentice-Hall, Englewood Cliffs, New Jersey, 1976.

68. Wismer, D. A. and Chattergy, R, "*Introduction to Nonlinear Optimization*", Wismer Associates, Inc., 1978.

69. Law, A. M. and Kelton, W. D., "*Simulation Modeling and Analysis*", McGraw-Hill, Inc., 1991, pp. 424-428, 485-503, 679-689.